

Pay for Performance in general practice - Quality to what price?

Kari Jussie Lønning



Master Thesis in health administration

Department of Health Management and Health Economics
Institute of Health and Society

Faculty of Medicine

University of Oslo

15.05.2013

Preface

What are the future solutions to meet the challenge of scarce health resources? We face a rapid increase in medical knowledge and technology, that leads to increased medical possibilities. The gap between what's possible and what our health budgets can cover, is increasing. The demand for health care seems to be forever mounting. Some of the most important reasons for increased demand are, to my opinion, that politicians as well as societies define more and more into the concept of health, and the demand driven health service (western world) instead of a health service based on medical needs. At the same time we have challenges by a demography change with the population getting older, living longer with more diseases and diagnoses. The search for financial and incentive systems that secure fair, equality and effective health services have been going on for many years and are highly alive – driven by an ambition to find the best system. Pay for performance (P4P) is an incentive system where the outcome of the treatment is rewarded. A highly invasive system, invading the patient-doctor encounter. Does it serve the mission of giving better health care services? I hope this thesis can contribute to illuminate this urgent question.

P4P is being used in many countries to a varying degree, both in primary care and secondary care (hospitals). My background is being a specialist in family medicine, and I had an interest in the use of P4P in primary care. UK is the country that has implemented the system nationally and most extensively, and I thus particularly wanted to study how GPs in the UK had experienced P4P.

I embarked MHA study at the University of Oslo, August 2011, and this thesis complete the master programme in health administration (MHA).

I have worked about 11 years as GP. For the last four years, I have been working at the Norwegian Medical Association as head of the department of primary care and psychiatry, later extended to head of Professional Development in the department of professional affairs. After embarking the study of MHA, I have been very engaged by the impact the incentive system Pay for performance, can have on the professional medical life as a doctor.

My supervisor has been Jan Frich. You have given me good and important advices during this process. I am deeply grateful. Thank you.

I will also give a huge thank to the GPs willing to take part in my interviews, they have contributed with important experience.

I will also give a thank to Anne Karin Lindahl, for giving me of her time to discuss the theme and sending me articles.

I applied the Fund of quality and patient security (Fond for kvalitet og pasientsikkerhet), and I got a scholarship to cover the expenses for the interviews, this was important contribution for accomplish the plan for this thesis.

I will dedicate this thesis to my dearly beloved father who died this Easter.

Summary

Aim and research question

The aim of this thesis was to explore P4P as incentive system in health care, and I wanted more specifically to do a review of the current literature, and also explore GPs' experiences with P4P and QOF in UK. My main research questions have been: Does P4P/QOF give quality improvement and to what price? How does QOF influence quality? What are the positive and negative effects? Does QOF influence ethical reflections or mindset of good doctoring?

Price is used in this thesis in a broad sense, by including both the price to pay in an ethical way, in relation to the patient -doctor relationship, professional, working force, money a.o.

Material and method

I have done unstructured literature searches, and collected data on GPs' experiences through interviews with five GPs in UK.

Results

The literature suggests that the QOF system seems to improve quality measures in diagnoses with clear endpoints, like diabetes, hypertension and some others. One could question if the improvement lasts, and also if improvement in quality measures causes better outcomes and care for patients. In sum, although there exists a huge amount of data, the evidence on the efficiency of P4P is scare and inconclusive. There are negative effects that have to be studied more thoroughly. In the interviews, GPs experienced that QOF had lifted "every doctor to the same bar" and given them a longed systematic approach to managing certain diagnoses. The system of QOF seemed to have eliminated the worst practices, but at the same time it left little room and no incentives for the exploration of excellence. GPs reported side effects of QOF that need to be taken into account.

Conclusion

In order to improve quality, we have to include all the dimensions of the term; Safety, Effective, Patient centered, Efficient, Timely, Equitable¹. Incentive systems have no value in themselves; they are just one among many tools to influence what happens in clinical practice. Future attempts to measure quality in primary care should take the unique complexity of primary care into account. We then need to engage multiple perspectives and multiple levels.

¹ http://www.peerpt.com/website/index.php?option=com_content&view=article&id=10&Itemid=10

Contents

Preface.....	2
Summary	4
1. Background.....	8
1.1 Historical perspectives on the health services in UK.....	9
1.2 Historical perspectives on the health services in Norway.....	10
1.3 The concept of quality in health care:.....	11
1.3.1 Structure, process and outcome	13
1.4 Aim for this thesis.....	15
2. Material and Method	16
2.1 Literature	16
2.2 Focus groups.....	16
2.3 Two interviews	18
2.4 Research Question.....	19
2.5 Analysis.....	19
3. The P4P – literature.....	20
3.1 Reviews on P4P, including other countries – what have we learned about P4P internationally ?	20
3.2 What have we learned about P4P? Single studies internationally	22
3.3 What are the doctors expectation of the P4P system?.....	23
3.4 Experiences from the UK from 2004 until today – different aspects of P4P	24
3.4.1 Blood pressure.....	24
3.4.2 Diabetes.....	24
3.4.3 Smoking and diabetes	25
3.4.4 Ethnic disparities	25
3.4.5 Socioeconomic disparities	25
3.4.6 Effects on mortality	26
3.4.7 Effects on access.....	26
3.4.8 Equity.....	27
3.4.9 Practice size	27
3.4.10 Practice organization.....	28
3.4.11 Doctors autonomy.....	28
3.4.12 What are doctor`s experience with P4P on their workday?	29
3.5 Different structural devises of QOF.....	30

3.5.1 Exclusion of patient from the QOF	
The literature points out the	30
3.5.2 Piloting the Quality indicators	30
3.6 Ethical aspects of the QOF.....	31
3.7 Does the QOF work – has it given the desired results?.....	33
3.8 Does P4P give value for money?	35
4. Results/Analysis of the interviews	36
4.1 Quality	37
4.2 Ethics - Ethical issues and the QOF.....	39
4.3 Positive effects of P4P	42
4.3.1 A more systematic approach.....	42
4.4 Negative effects of P4P:	44
4.4.1 Doctors versus patient centred care	46
4.4.2 Overtreatment:.....	47
4.4.3 Political tool.....	49
5. Discussion	52
5.1 The main findings	52
Positive effects:	52
Quality	52
Negative effects:.....	52
Ethical issues	53
5.2 Methodological considerations.....	53
5.2.1 The two interviews	54
5.2.2 Reflexivity	55
5.3 Thematic discussion:	56
5.3.1 A more systematic approach?	56
5.3.2. Patient centred health care.....	57
5.3.3 Overtreatment.....	58
5.3.4 Political tool.....	58
5.3.5 Quality	60
5.3.6 Ethics.....	64
6. Conclusion	67
7. Reference list:.....	69

;

8. Appendix.....	74
Appendix 1.....	74
Appendix 2.....	75
Appendix 3.....	76

1. Background

My interest and curiosity for Pay for Performance (P4P) was first awakened in 2008 after speaking to a General Practitioner (GP) from United Kingdom (UK) at a conference in Oslo. I was lucky to join her for a separate lunch and got the opportunity to discuss and learn about the system of P4P or Quality and Outcome Framework (QOF). She presented me statements like: “P4P has ruined our profession”, “It’s been the death of reflective medicine”. “You have to fight against it – do not let it ruin your solid family medicine in Norway.” We were several GPs at this “separate lunch” – and also some representing the Norwegian authorities, so it certainly set of a discussion.

Without any further knowledge about P4P, I joined the conference “Helse i Utvikling 2010”. This is a conference arranged every second year by the Knowledge centre for health services in Norway. I heard Tim Stokes from NICE, UK, present “Development of guidelines and quality indicators in primary care”. After the presentation I asked myself: What kind of results must be demanded to justify such a huge arrangement as he presented with the QOF indicators and guidelines? What about the transaction costs? Not only with regards to money, but also in terms of the demands put on the working force. A whole department dedicated to develop indicators, several departments to develop guidelines, and a huge group of controllers to see that the system work. I was really not less sceptical after this presentation.

In autumn 2011, I embarked on a master programme in health administration, and I got the opportunity to discuss and learn different opinions both from professors and fellow students on P4P as a system. I then realized that this system was, to some of the health economists, really like a dream - a dream of scientific challenges and opportunities to get lots of interesting questions into the academic of Norwegian health economics. I got worries on behalf of the future of general practise and family medicine in Norway, although it should be mentioned that there were some scepticism among the professors in health economy as well.

On the 1 of September 2011, the Norwegian Directorate of Health arranged a Policy seminar with the title “Introducing Pay for Performance (P4P) in the Norwegian health system scoping relevant international experiences”. This confirmed my concern about getting P4P or QOF introduced in Norway.

I just had to learn more about P4P - and as my background is being a specialist in family medicine - I wanted to learn more about P4P in general practice/primary care.

After a discussion with my supervisor, I decided to do a study into P4P and QOF, investigating the literature and explore experiences of GPs in the UK.

1.1 Historical perspectives on the health services in UK

Core features of primary care in UK have been constant since the National Health Service (NHS) was established in 1948. There is universal registration with a single practice of the patient's choice, and medical services in the primary health services is provided by general practitioners (GPs)(1). In 1990 a modest P4P was introduced in primary care in the UK, in the form of payments for reaching target levels of childhood immunization and cervical screening (cytology). This led to increased performance within these targets, followed by a slower reduction in socioeconomic inequalities (2). In 1998, the NHS embarked on a widespread program of quality improvement under the general heading of “clinical governance.” (1) This included the development of national clinical guidelines and national service frameworks to guide implementation of improvement activity; a body to make recommendations about cost-effective treatments in England (NICE, www.nice.nhs.uk); the introduction of annual appraisal for all NHS doctors; district-wide audits of clinical care, with identifiable data being shared with practices and sometimes with patients; and a range of local financial incentives schemes for quality improvement. These were associated with significant improvements in quality of care. (1) Although these improvements gave an increased quality in health delivery, there was given two reasons for changing the incentive system in 2004; international studies indicating to low quality in UK health care and the political statement saying England did spend less money in health care in comparison with other Western countries (3). So in 2004, a new and quite ambitious scheme – referred to as “a radical experiment” by some (4) or the boldest proposal of its kind ever attempted anywhere (5) –Quality and Outcome Framework (QOF) a Pay for Performance (P4P) scheme was introduced in general practice. The NHS committed 1.8 billion pounds in additional funding over a period of three years for the new pay-for-performance program for GPs (6). This performance related pay arrangement included 146 quality indicators covering management of chronic diseases (ten different diseases), practise organisation, patients experience with respect to care and extra service. The

management of chronic diseases was divided into sub groups each with targets to be reached, and the goal was to reach fulfillment each target.

The practise organisation was scored, together with the patients surveys and the practise`s plans for following up patients surveys (7). The maximum score was 1000 points/ year. Every point gave an amount of pounds, in 2004-2005 it amounted 76 pounds and in 2006 it was raised to 125 pounds per point. The first year the GPs reached a score of 84% (6,7) and 96,7 % for the clinical indicators(6). This gave a raise in salary for about 20-25% (6) This greatly exceeded the 75 % predicted when the scheme was negotiated and consequently the cost to the taxpayers was considerably more than expected. Since 2004, new clinical areas have been introduced and payment thresholds have been raised gradually. In 2006, the indicator set was revised; all minimum and some maximum payment thresholds have been raised, 30 indicators have been left out or modified, and 18 new indicators have been introduced (6), in 2006 the indicators counted 136 and in 2011/2012 it was 134 indicators (8). An important feature of QOF is that GPs can exclude patients if they judge that incentivized care would be inappropriate for particular individuals (9,10). A scheme to tie GP payments directly to patient experience survey scores was introduced in 2008 (11), but it proved problematic (12), and was withdrawn in 2011.

1.2 Historical perspectives on the health services in Norway

Despite many differences, the Norwegian and the English health system have many historical similarities. General practice was introduced as a medical specialty in 1985. Although young, the family medicine is now a solid, proud and well established specialty in Norway and have a good reputation also beyond Norway`s borders.

It was not always like this. In the 1960 a group of GPs started to work for family medicine as an independent academic discipline" at the universities in Norway. This was the first step to get family medicine away from a "second hand area in medicine". There were problems getting young doctor or doctors at all, to work in primary care. When the two Universities in Norway at that time, started to approve family medicine and teaching it as a subject in the undergraduate medical curriculum and there were financial arrangement supporting the family medicine (1963-67), it slowly got a higher status. The speciality in family medicine is the only one in Norway with mandatory recertification every fifth year.

The educational content of the specialty of family medicine is constantly being revidated and controlled by the Specialty Committees (spesialitetskomite)². In Norway, there are 44 different medical specialties, everyone having a Specialty Committee. The Specialty Council³ (Spesialitetsrådet) are supervising all the 44 Specialty Committees. The Specialty Council monitors specialist training and deals with matters of principle that arise in that connection. Both the committees and the Council act as expert advisory bodies to the Board of the Norwegian Medical Association and other bodies on matters relating to post graduate- and continuous education. They are considering the continuous medical education and the continuous professional development in the specialist education and continuing education. In 2001 we got a reform in Norway- The regular GP scheme (Fastlegereformen). Every inhabitant in Norway hold the right and possibility to get listed with a GP. The financial sources giving GPs in Norway their salary is threefold. The payment is both from the patient, about 1/3 (regulated tariff by the government), about 1/3 from the authorities (fee for service) and about 1/3 from the community (per capita.) Most of the GPs are self-employed (about 95-97%) only about 3-5 % are employees (community). As self-employed they are strictly regulated by a tariff (normal tariffen) set in the annually bargaining by the authorities and the Norwegian Medical Association.

1.3 The concept of quality in health care:

The term Quality in health is given many definitions and descriptions. The NHS has given the following description and definition;

"Quality' is a term used with different meaning within the NHS and covers many aspects of service provision, including waiting times for treatment, convenience and accessibility, cleanliness of facilities, and patient involvement, as well as the quality and effectiveness of clinical care."

In the Next Stage Review (NSR 2008), quality is defined by Lord Darzi as care which is; "Clinically effective, personal (patient experience) and safe".

Sir Ian Carruthers, Chairman of NHS South West, has expanded on this theme:

² For each of the 44 specialties in Norway there is a five-member Specialty Committee, which deals among other things with the rules and the contents of specialist training.

³ The specialty Council has 16 members, they represent the faculties of Medicine, Board of health, The Directorate for Health, The Norwegian patient organisation, The association of Local Authorities and various units in the Medical Association's organization.

“I would define quality as being safety, the experience of the individual, evidence-based best practice, access and taxpayer value, all of which string together to say: how do we improve treatment and the quality of life of individuals?”⁴

The Institute of Medicine (IOM), an independent organisation in USA, defines healthcare quality as «the extent to which health services provided to individuals and patient populations improve desired health outcomes.» (13). The IOM has designed six “aims” for improving the delivery of care in the United States. PeerPoint keeps these aims in mind when creating unique programming with a goal of improving patient outcomes.

- **Safe** - Avoiding preventable injuries, reducing medical errors,
- **Effective** - Providing services based on scientific knowledge (clinical guidelines),
- **Patient centered**- Care that is respectful and responsive to individuals
- **Efficient** - Avoiding wasting time and other resources
- **Timely**- Reducing wait times, improving the practice flow and
- **Equitable**- Consistent care regardless of patient characteristics and demographics⁵

The Norwegian health authorities also defines quality through six aims, in both “National strategy for Quality improvement in Health and Social Services” (*Nasjonal strategi for kvalitetsforbedring i sosial- og helsetjenesten... Og bedre skal det bli*) and the National Healthplans:

- **Effective** – give improved health (virkningsfulle, føre til en helsegevinst)
- **Safe and secure** - (trygge og sikre, unngå uønskede hendelser)
- **Involves the patients** – give the patients influence
- **Coordinated** – coordinate to give the health care continuity
- **Efficient**-avoiding wasting time and other resources
- **Available and fair** – (er tilgjengelige og rettferdig fordelt)⁶

The Norwegian Medical Association (NMA) defines quality to be related to the result of patient treatment, regarding both the patient experience and the medical diagnosis and medical treatment⁷.

⁴ <http://www.publications.parliament.uk/pa/cm200809/cmselect/cmhealth/53/5307.htm>

⁵ http://www.peerpt.com/website/index.php?option=com_content&view=article&id=10&Itemid=10

⁶ <http://www.regjeringen.no/nb/dep/hod/dok/regpubl/prop/2010-2011/prop-91-l-20102011/21.html?id=639058>

The definition presented above give a fairly wide scope, and I think that is important and necessary when trying to fill this complex concept – quality in health. As there are many definitions and descriptions of quality, there are also many measures introduced to improve the quality of care. One of them is Evidence Based Medicine (EBM). The increasing availability of medical evidence in clinical practice was expected to improve the quality of care. However this has not been realised (14), and an impressive number of studies conclude that implementation of guidelines is difficult (14). Donabedian used the triangle of structure, process and outcome and this gives a good frame for identifying three types of evidence into the quality reflection; medical, contextual and policy evidence as presented in the article from Van Driel et al (14). The knowledge base of medicine is increasing, but it has major flaws and gaps hempering its applicability in primary care. Why is that? In Norway we have a huge challenge due to almost all research done in the hospitals or specialist care, on a patient population quite different from the population in the primary care. This is also a challenge in other countries, as far as I have understood. Many medical questions, in clinical practice and specifically in primary care, have not been studied (14).

Quality of care must always be seen in the context of the health care system. This is important in the evaluation of P4P and the other financial system – like Fee For Service (FFS) and capitation among others. Communication is the most important aspect of context, and evidence shows that more important than the length of the consultations is the quality of the doctor- patient relationship. This relationship is also important in aspect to adherence to therapy and even outcome (14). How are both communication and the relationship affected by financial incentive system? The policy evidence takes in efficiency and equity. Due to scarce healthresources the importance of economic analysis are increasing and are therefore taken into the planning for health care intervention and innovations. It's not difficult to see the need of financial analysis, but there is a risk getting a one dimensional focus. In my opinion there must be more than solely economical focus in to the calculation, as for example ethical, equity and science assessments.

1.3.1 Structure, process and outcome

As we seek to define quality, we soon come aware of the fact that several ways of defining the term are both possible and legitimate, depending on where we are located in the system of care and on what the nature and extent of our responsibilities are (15). Donabedian has described circles in layers to draw the levels at which quality may be assessed. The inner core

⁷ <https://legeforeningen.no/PageFiles/26457/Med%20kvalitet%20som%20ledestjerne.pdf>

of the circle or the first level, is the health personals performance, next layer or level is the amenities and then care implemented by patient, and the in outer circle care received by Community is described (15, fig 1). This picture shows how “everything depends on each other” and the quality in health care systems will never get better than the weakest link in the chain. And before going into the determinators for quality process, structure and outcome, I will quote Donabedian on his statement on the mainstay in performance: “Clearly, the interpersonal process is the vehicle by which technical care is implemented and on which its success depends.” (15).

Structure is an important determinator for quality, and is a premise for the interaction between the health care system, society and the individuals in society. Structure refers to the setting in which care is delivered including adequate facilities and equipment, qualification of care providers, administration structure and operations of programs. Using this approach, good care settings and supporting structures contribute to good care. Structure variables are often concrete and accessible, making them relatively easy to assess.

Process determine quality and is tight linked to the structure by continuous interaction. Process is the interventions and interactions between the patient and health care provider. Communication skills in both patient and health care worker are extremely important to get a good process as mentioned earlier. The medical decision making process is supported by guidelines and protocols and so on. EBM is an important contribution as are also internet with free asses to medical scientific bases and decision support tools.

Outcome indicators can be symptoms and complaints from the patient and medical parameters, quality of daily life indicators, patient satisfaction and social equity. There is a shift in focus from problem oriented to goal oriented(16).

Back to the medical evidence – the body of evidence is increasing rapidly and the evidence is being more available much due to the emphasis on EBM and the way EBM is made available through digital Electronic Medical Journals (EMJ), support tools for the EMJ and internet. An important restriction of the practical usefulness of medical evidence is that most research is not driven by clinical relevance, but rather by commercial interest (14) As an example there are many studies and vast number of publications on pharmacological treatment, but only a few studies are published on the effect of exercise and lifestyle changes(14). In my opinion this also give reason for concern and question must be asked to what extend the research dimension of EBMs is biased? And there is also a danger that EBM pursues what is possible,

rather than what is relevant(14). The unequal initiatives in research respectively in primary- and specialist care are also a big challenge. Most of the evidence based knowledge are based on a population from selected patient population in specialist care.

Just as a clear and uncontested definition of quality in primary care are unrealistic there are some specific component that are widely agreed on as central to the idea of quality(17).

Barbra Starfield has described them as four unique features of a primary care service; access, personfocused care over time, comprehensiveness and coordination. Another important aspect is continuity of care (17), cf. chapter 5.

1.4 Aim for this thesis

I wanted to know - and to learn more about P4P and QOF. I wanted to know the arrangement and history of the framework. Get to know the effects, positive and negative of QOF on different aspects of medicine. Get to know the results measured and evaluated so far. I will present my research questions later – in chapter 2. The aim of this thesis is to explore P4P as incentive system in health care, and I will more specifically review of the current literature, and also explore GPs' experiences with P4P in UK.

Why was (or is) this important to me? One reason is obvious; I had to challenge my scepticism towards P4P and see if it still was the same after I had attained more knowledge. But behind my scepticism there are several different questions. And these several other questions aroused my curiosity to and the need for getting to know more about P4P.

I just had to learn more about P4P and QOF and the aim of this thesis to get knowledge to see if it make me able to give some thoughts or assessments about this incentive system.

2. Material and Method

2.1 Literature

During the period of January – June 2012 I did non-systematic searches in PubMed, Medline, Best Practice (some), Up to date (some) and Google.com/Google.no to find material about P4P. I used a form to organise the key words (Appendix 1). I have also reviewed reference lists of several of the articles. Searches were done with help from a librarian at library of Faculty of medicine University in Oslo, by myself and I have also been tipped about some articles from colleges.

I found over 1880 hits on P4P and quality (03.01.2012) in PubMed and other searches gave an enormous number of hits. So, I had to make a selection among the articles and studies. I have tried to search for articles who evaluate different dimensions of QOF and P4P with the perspective that tempted me into the choice of this theme. I am sure there are many more relevant articles, but I have selected a sample of the most relevant articles which I cite in this thesis.

The studies and articles were analysed according to the subject, theme or process they described. The articles are presented in chapter 3.

In addition to studying the literature, I wanted to get an insight into experiences with P4P/QOF by talking to colleges, GPs in UK, about their experience with P4P or QOF as they have been working in and with the system for about nine years.

2.2 Focus groups

My plan was to conduct two focus groups, with the hope of recruiting five to eight GPs in each group.

Sampling theory

I had given different Qualitative sampling methods for interviews and focus groups, much thoughts before I started to recruit participants (18). I had considered "typical case sampling" – trying to get together the average GPs, considering geography, patient population and opinion about QOF. This would have presupposed that the literature, and the previous research could give me information about characteristics leading me to these GPs. To my knowledge there is no such information, and the "typical" GP with regards to P4P would be difficult to identify. The same would be the case for "critical case sampling". If there were to

find theoretical models or previous research who could lead me to especially information-rich and thus particularly illuminating participants.

I was thinking about “deviant case sampling” - sampling the most extreme opinions. A thought was to have one group with GPs really pro QOF and one focus group with GPs negative towards QOF. Another thought was to gather GPs with “maximum-variation”. That would have given as wide range of perspectives as possible and given the opportunity to capture the broadest set of information and experiences. The sample would include “typical-“, “deviant-“, and “critical cases”, plus any other perspectives identified. Again my lack of knowledge about UK GPs from the inside, did not give me the opportunity to collect or recruit under this goal.

Another sampling theory for the focus group interviews is “confirming-disconfirming sampling” - Sampling both individuals or texts whose perspectives are likely to confirm the researcher’s developing understanding of the phenomenon under study and those whose perspectives are likely to challenge that understanding (18). The sample would include GPs whose experiences would likely either confirm or disconfirm what I already had learnt.

My method for recruitment is to be described by the category of “snowball sampling”. This is because several of my participants were recruited through asking current participants to recommend others.

Sampling the GPs

I contacted Royal College of General Practitioners (RCGP) in UK to inquire if they could help me getting in contact with GPs. I first sent a request in April 2012, and I received a very positive response. But after months with emailing and reminding, RCGP did not send out an invitation before late October/early November 2012. I did not get to know who’s addresses, how many, or based on what criteria (other than GPs in London) they choose the addresses for the invitation. One GP signed up after this invitation. I contacted colleges in Norway who had contacts in UK, and by them I got some email addresses to persons I could contact. The people I contacted were professors and teachers, so I asked them if they could help me to get in contact with GPs. They gave me the advice to contact RCGP.

I then started to use internet for recruiting. I searched randomly and found some professors and teachers at GP institutes in London and through one of them I got recruited another GP. Then this GP recommended me to join a group on LinkedIn (to give the participants anonymity I wan’t name the group here), and he recruited his girlfriend (working in another

city and felt free about their opinions of QOF). I announced the invitation to the interview in this group and got one more participants. I searched for GP, London in LinkedIn and started to send invitations direct to the GPs I found in London and this way I got two more GPs totally five GPs.

2.3 Two interviews

The two focus groups had to be moderated into one depth interview and one mini-focus group interview.

The two interviews took place on the 6th of December, in London in a hired location at St. Pancras station. The location was found through a contact in RCGP. The criteria I gave my contact, was that it should be a quiet and nice place. Both criteria were met at St. Pancras, Searcys location.

I made an interview guide, and got input from my supervisor (Appendix 2). Together we decided that I should use the guide not as a strictly interview tool, but as a reminder to secure that the interviews were guided into the different themes.

The participants was handed a letter with information about the project before the interviews started where I guaranteed them anonymity (Appendix 3). They also recieved a gift card from a book store, as a small reward for participating (50£).

The first interview took place between 8 am – 9 am. The interview was with GP who had worked for several decades in the same practise. He worked in London. The area was described as a mixture of suburban and city characteristic. The interview lasted for 47 minutes. Since this interview was with one participant, I prepared for it as a non-structured depth interview. I wanted the participant to reflect freely about his experiences and thoughts about QOF. And therefore, as described earlier, I used the interview guide as a reminder and not strictly as a guide. I tried to ask open questions and facilitated a relaxed atmosphere. I wanted to learn from his “world of experience” as Tjora describes in *Kvalitative forskningsmetoder i praksis* (19).

The second interview took place between 10-11.30 am. This interview was with four younger GPs (A, AA, E and L). This group can be described as a mini-focus group (19). Three of them qualified as GPs in 2008, the fourth qualified in August 2012. Two worked in London, for most of the time in the North West part, in practises varying from 4000 – 12 000 patients. One had experience both from London and another small town (urban to very rural) and the last

one worked in a “tiny little town”... on the south east coast in UK. They all worked as locums, salaried or freelance. The interview lasted for 1 hour and 14 minutes.

In the mini-focus group I also used the interview guide freely and asked the questions open and wanted to facilitate an open and free conversation, just making sure that I got them to reflect on some themes that were essential for the study.

Both interviews were relaxed and floating conversation in my opinion, in a good and relaxed atmosphere. Both interviews were recorded by a digital recorder and later fully transcribed.

I have also talked to research colleges in Norway about their knowledge about P4P, to learn more about different views.

2.4 Research Question

My research questions have been:

Does P4P/QOF give quality improvement and to what price? How does QOF influence quality? What are the positive effects of the framework? What are the negative effects? Does the QOF influence the ethical reflections or the mindset of good doctoring?

2.5 Analysis

The transcribed text was analysed. I coded statements and themes depending on whether they were describing positive or negative effects of the QOF, and made subcodes in each of these two categories depending on different themes or topics addressed. In both interviews the participants highlighted the positive systematic effects QOF had given them. All participants also addressed many negative effects like the theme health care focus: Who is put in the centre the patient or the system/doctor? Overtreatment was an issue addressed in both interviews and QOF as a political tool was also addressed and was an important issue for the participants both in the depth interview and in the mini-focus group. Two categories, quality and ethics, were coded as main categories and statements giving information to the two categories, were gone through separately.

It was the first time I did a mini-focus group interview, and first time performing interviews in a foreign language. My native language is Norwegian, and I started both the interviews by asking the participants to let me know if I my language was unclear or asked for permission to stop them if I didn't understand them. The interviews were conducted without any language problems.

3. The P4P – literature

I have organized the literature under the following headings:

What have we learned about P4P internationally? What have we learned about P4P – single studies internationally? What are the doctors expectation of the P4P system? Experiences from the UK from 2004 until today – different aspects of P4P, Different structural devises of QOF, Ethical aspects of the QOF, Does the QOF work – has it given the desired results? Gives the P4P value for money ?

3.1 Reviews on P4P, including other countries – what have we learned about P4P internationally ?

Through my search in literature on quality and primary care, I found that over the last 20 years many quality-improvement initiatives have been tried in the UK and in many other countries. The initiatives are varying from education, to clinical audit and guidelines (4). There is a growing interest in different economic incentive systems like paying for performance, as a means to align the incentives of health workers and health providers with public health goals(20). This is seen in many countries, by many funders and governments. Even so, there is to note that there is currently a lack of rigorous evidence on the effectiveness of these strategies in improving health care and health (20,21) This is particularly underlined for low – and middle -income countries (20), but found in many studies worldwide (15). Moreover, paying for performance is a complex intervention with uncertain benefits and potential harms, if I am to use a statement from the review from Witter et al 2012 (20). In another review by Van Herk et al 2010 (22) they summarizes evidence, obtained from studies published between January 1990 and July 2009, concerning P4P effects, as well as evidence on the impact of design choices and contextual mediators on these effects. In the 128 studies included they found wide range of effects of P4P from negative to absent, to positive or very positive depending of the target and program. In this context negative means less quality improvement compared to non P4P use and not a quality decline. They found for preventive care more conflicting results for screening targets than immunization targets. Acute care most frequently failed to be affected by P4P across the studies in the review (22). In chronic care, diabetes was the condition with the highest rates of quality improvement due to P4P implementation. Positive results were also reported for asthma and smoking cessation. This contrast with finding no effect with regards to coronary heart disease (CDH) care (22). The

review states that throughout the studies included it seems like P4P have contributed to closing gap for performance differences and continues: depending of the design choice process indicators generally yielded higher improvement rates than outcome measures, with intermediate outcome yielding in-between rates. A review by Emmert et al (21), based on articles with full economic evaluations, and partial economic evaluations considering costs and consequences of the P4P intervention simultaneously were studied. The review included nine studies, three full economic evaluations and six studies were classified as partial economic evaluations. The full economic evaluations, could not demonstrate efficiency of P4P. The partial economic evaluations showed mixed results, but several flaws limit their significance (21). Emmert et al. Concludes the results show that evidence on the efficiency of P4P is scarce and inconclusive. P4P efficiency could not be demonstrated (21).

Furthermore studies reporting involvement of stakeholders in target selection and definition seem to have found more positive P4P effects. The P4P positive incentives (financial rewards) seem to have generated more positive effects than incentives based on a competitive approach (22). The positive effect was higher for initially low performers compared to already high performers (22). The review summarize six recommendations who are supported by evidence throughout the 128 studies included:

- Select and define P4P targets based on baseline room for improvement.
- Make use of process and (intermediary) outcome indicators as target measures.
- Involve stakeholders and communicate the program thoroughly and directly throughout development, implementation, and evaluation.
- Implement a uniform P4P design across payers.
- Focus on quality improvement and achievement.
- Distribute incentives at the individual level and/or at the team level.

The review concludes; it is need for more research and these should address the issues where evidence is absent or conflicting (22). This goes along with the authors conclusion in the review by Witter et al also; «The current evidence base is too weak to draw general conclusions; more robust and also comprehensive studies are needed. Performance-based funding is not a uniform intervention, but rather a range of approaches. Its effects depend on the interaction of several variables, including the design of the intervention, the amount of additional funding, other ancillary components such as technical support, and contextual factors, including the organizational context in which to be implemented»(20).

3.2 What have we learned about P4P? Single studies internationally

There is an increasing amount of studies and literature on the subject P4P and financial incentive systems in health care, especially the last two years. There is increasing international interest in pay for performance in primary care. The introduction of these performance system reflects concerns about three interlinked issues – the variation in performance and quality, the emphasis on driving improvements in performance and ensuring high quality in primary care (17). Performance as a measure of quality depends upon what the performance standards are and how they are measured. This is very important in relation to P4P schemes. Another important premise for evaluating performance as a measure for quality is what we define into quality (cf. chapter 1.3,4 and 5 Quality). There are many definitions – and quality differs in the frame of context. The literature is rich in articles about this issue from many different countries (23-27). And many countries compare their system to the UK system of QOF. In one study by McDonald et al (23) 40 primary care physicians in California and England, 20 each, were interviewed and compared about unintended consequences. The conclusion was that the design and implementation of the incentive programs are related to the extent of unintended consequences. There were more unintended consequences in England than in California (23). Crooson et al compared (26) the quality of chronic illness care in US and UK before implementing P4P. They analyzed medical records data collected before QOF implementation from randomly selected patients with diabetes or coronary artery disease in 42 UK and 55 US family medicine practices. The conclusion was that there were given more standardized care in UK than in US, but no better intermediate outcomes. Crosson et al (26) emphasises the investment done in UK prior to the QOF and especially into building the capacity of primary care practices to effectively use electronic medical records (EMR). They stresses the need for this investment in US primary care and say that maybe P4P is a way to go, but not without adequate preparations: »pushing P4P into practice without adequate preparation could lead to premature rejection of this policy tool and another missed opportunity to reorient the US health care system towards more effective care coordination and preventive delivery» (26). Bell and Levinson from Canada (24) reflects on the question whether P4P incentives result in improved patient outcome or do they just lead to improved documentation and use of processes of care? They go on reflecting «No physician chooses to provide low-quality care,...» But they call for the Canadian physicians to have access to quality information technology resources and measurement system to inform physicians about

health care quality and clinical outcomes. They would like to receive individualised feedback and suggested intervention strategies to be helped to better pursue excellence in patient care. Pride and professional ethos have been identified as drivers in quality improvement, they reflect. They agree with rewarding quality care that improves outcomes, but what is the best method still remains uncertain (24). A study from France by Chauvel et al. (25) looked into what characterised the doctors not performing optimal in relation to prescribing HbA1C for patient with diabetes. They studied 2545 doctors and 41453 patients. There was a large variability between GPs, also after adjusting for patient characteristic. Doctors who were female, young, working in a group practice, participating in quality-control groups, and who had lower patient load prescribed the three or four recommended tests more often. There may be other ways to improve quality than incentive systems alone? By facilitate doctors training and doctors possibility to practice in groups? (25). Kirschner et al (28) studied 65 general practices in the south of the Netherlands in an observational study with a pre- and post-measurement (Before and after introducing a P4P program). A P4P program designed by target users containing indicators for chronic care, prevention, practice management and patient experience where introduced. After one year significant improvement was shown for the process indicators for all chronic condition (+7,9 – 11,5%). Five outcome indicators significant improved as well as patient`s experience with GPs functioning and the organization of care.

3.3 What are the doctors expectation of the P4P system?

In a study asking 1668 randomly selected general internist in USA about their view on P4P, Casaline et al. found a strong potential support for financial incentives for quality, but less support for public reporting. Large majorities of respondents stated that these programs will result in physicians avoiding high-risk patients and will divert attention from important types of care for which quality is not measured (29). They conclude by stating that evaluation of P4P and public reporting programs should be explicit designed to assess possible unintended consequences in disparities in health care delivery, on physicians who practice in areas of low socioeconomic status, on the quality of care in important areas and on the quality of care in important areas of physicians practice not included in the program being evaluated.

3.4 Experiences from the UK from 2004 until today – different aspects of P4P

As UK have the most ambitious concept of pay for performance and as I said earlier a Health insurance system much like ours in Norway, I have focused on UK and articles evaluating the UK QOF/ P4P system.

The clinical diagnoses that are being evaluated in this research are diabetes (26,30,31,32,33,34), hypertension (26,27,35,36,37), asthma (34,38) and chronic heart deceases (CHD)(20,34,37). Also smoking are subject to many studies (33). These are all pretty clear clinical measurable diagnoses or health behavior, where improved quality can more easily be defined as improving the measures (lover Blood pressure, lover HbA1c and so on). Many of the studies from UK have quite a huge data collection. The national database for registration in the «tic boxes» (P4P schemes) in the GPs EMR (over 8500 practices (27)) gives a really amazing research material and give a solid basis for evidence.

3.4.1 Blood pressure

Studying the impact of the P4P on blood pressure monitoring (27,35,36,37), the results are fairly clear showing that the blood pressure monitoring and control, may have improved. But also showing that there is a need to follow up patients who are older and more socioeconomically deprived once they are diagnosed, as well as prescribing antihypertensive therapy to younger patients, who are likely to benefit from early intervention (27). In Serumaga et al`s study(36) of 470 725 patients with hypertension diagnosed between January 2000 and August 2007, they found that good quality of care for hypertension was stable or improving before pay for performance was introduced. Pay for performance had no discernible effects on processes of care or on hypertension related clinical outcomes. Generous financial incentives, as designed in the UK pay for performance policy, may not be sufficient to improve quality of care and outcomes for hypertension and other common chronic conditions(36).

3.4.2 Diabetes

Diabetes is thoroughly studied (30,31,32,33,34) under the P4P incentives. The conclusions are not clear due to lack of control practices (32,34) in whether it is P4P incentives that caused the observed increasing achievement of targets and reducing problems of low performance, or

if other development efforts may have been influential (32,33). «There was already evidence of improving quality of care before introduction of QOF. The QOF targets are designed for audit rather than best practice, and practitioners may be utilizing clinical practice guidelines that recommended more stringent targets »(32). Campbell et al (33) found that although these questions must be raised, there was a modest acceleration in improvement for two of these three conditions they studied: diabetes and asthma.

3.4.3 Smoking and diabetes

Have the P4P improved smoking cessation in diabetic patients? This was studied by Millett et al (33) and they compared data from 2003 (before P4P was introduced) and 2005-2006 (after introduction of P4P) in almost 4300 patients. They found that there was a association with reduction in smoking prevalence among patients with diabetes in primary care settings. They recommend health care planners in other countries to consider introducing similar incentive scheme for primary care physicians as in England (33).

3.4.4 Ethnic disparities

The impact of P4P on ethnic disparities are studied in several articles (30,31,37) in two of the articles taken into this report, ethnic disparities is looked at for diabetes care. The conclusions are clear; even though the introduction of P4P was associated with improvements in the intermediate outcome of diabetes care for all ethnic groups, the magnitude of improvement appeared to differ between ethnic groups (30). Policymakers should consider the potential impact of P4P incentives on health disparities (like disparities in mortality from cardiovascular disease and the other major complications of diabetes) when designing and evaluating such programs (30,31). Ethnic disparities are also to be considered in CHD treatment and specific targets may also be needed (37).

3.4.5 Socioeconomic disparities

The achievement gap between least and most deprived areas have been studied (35,39,40). Ashworth et al found by obtaining data from 8515 practices in 2005 (year one) and 8264 in 2007(year two) that improvements in achievement of monitoring blood pressure have been accompanied by the near disappearance of the achievement gap between least and most

deprived areas(35). In a US study by Chien et al(39) they found that physician organizations performance score vary by the socioeconomic status areas and the P4P programs have to account for this. They had 160 physician organizations included in their study. Doran et al (40) examined the relation between socioeconomic inequalities and delivered quality of clinical care in the first three years of P4P scheme in England. They found in analyzing automatically extracted data from 7637 general practices in England, that increase in achievement during the time was inversely associated with practice performance in previous years, but was not associated with area deprivation. They conclude that their results suggest that financial incentives schemes have the potential to make a substantial contribution to reduction of inequalities in the delivery of clinical care related to area deprivation(40).

3.4.6 Effects on mortality

Have the program reduced population mortality? This was studied by Fleetcroft et al(41) by combining twenty-five clinical quality indicators with controlled trial evidence of mortality benefit and condition prevalence and the difference in performance before and after program implementation, to estimate the potential mortality reduction per indicator. Improvement was adjusted for pre-existing trends where data were available. The results showed that P4P program potentially reduced mortality by 11 lives per 100 000 people (lower-upper estimates 7-16) over one year, as performance improved from baseline to the target for full incentive payment. If all eligible patients were treated, over and above the target, 56 (29-81) lives per 100 000 might have been saved. For the 2006 contract, mortality reduction was effectively zero, because new baseline performance for a typical practice had already exceeded the target performance for full payment (41). The program may have delivered substantial health gain, but potential health gain was limited by performance targets for full payment being set lower than typical baseline performance. Information on both baseline performance and population health gain should inform decisions about future selection of indicators for pay-for-performance schemes, and the level of performance at which full payment is triggered (41).

3.4.7 Effects on access

As in Norway and many other countries, England gives improvement of access to general practice priority. Improved access and more rapid access to care was one of the main causes to the reform in 2004(42-44). In 2006/2007 an annual national survey of patient experience of access was introduced, with financial incentives to practices based on the findings of the survey among their own patients (44). In 2008 a separate voluntary scheme for extended

hours was introduced. This enables practices to earn additional income by offering appointment outside the core contracted hours of 8am to 6.30 pm on weekdays(43). Campbell et al studied 42 representative practices by sending questionnaire to serial samples of participants of patients (42). They found a modest improvement in access to care for patients with chronic illness, but all patients did find it somewhat harder to obtain continuity of care. This can be related to the incentive to provide rapid appointments or to the increased number and specialized clinics in primary care (42). Addink et al studied results from the access survey in 222 general practices in 2006/07 and 2007/08 comparing the results. They found that the launch of the incentive scheme related to the access survey was not followed by convincing improvements in patient experience of access (44). The effects of the introduction of separate scheme for extended hours were studied by Morgan and Beerstecher (43). The terms and conditions of the extended hour scheme are loosely implemented and this may have limited the apparent effectiveness of the scheme but the conclusion was that demand for additional opening in primary care is only influenced by Saturday appointments. Satisfaction with opening hours responds to increased capacity, but is not linked to a specific time period (43).

3.4.8 Equity

Has the P4P improved the equity in health care in the UK? Boeckxstaens et al (45) studied this by performing a systematic literature search. They identified 317 studies, but excluded 290. None of the studies assessed equity in access to health care, but had looked into equity in treatment and (intermediate) treatment outcomes. They found that overall quality scores generally improved, and for the majority of the observed indicators, all citizens benefit from this improvement. But they also found that the extent to which different patient groups benefit tends to vary and to be highly dependent on the type and complexity of the indicator(s) studied. In general, the introduction of QOF was favorable for the aged and males (45). They conclude by saying that evaluating P4P initiatives in a broader health systems impact assessment strategy with equity as a full assessment criterion is of utmost importance (45).

3.4.9 Practice size

Is there any differences between practices of different sizes and performance under the QOF/P4P program? Both Vamos et al (46) and Doran et al (47) studied this and they have much the same conclusion or interpretation. Doran et al (47) found by a longitudinal analysis

of quality scores for 48 clinical activities in 7500 family practices, no evidence that size of practice is associated with the quality of diabetes management in primary care. Vamos et al enrolled 422 general practices in a retrospective open-cohort study and conclude; «The P4P programs appear to benefit both large and small practices to similar extent» (46). Doran et al (47) concludes that «the effect of the P4P scheme appears to have been to reduce variation in performance, and to reduce the difference between large and small practices».

3.4.10 Practice organization

Have the P4P program led to changes in the practice organizations? Checkland et al (48), Alyahya (49) and Maisey (50) have looked into this. Checkland et al (48) used interviews and observation to investigate in depth the impact of the QOF in four general medical practices. Alyahya (49) based his study on semi structured interviews with four GP practices in UK, involving 39 informants. Maisey et al (50) performed qualitative semi structured interview on 24 clinicians; 12 GPs and 12 nurses. They all observed changes, in practice organizational structures; an increased role of information technology; a move towards a more biomedical form of medical care; and changes in roles and relationships, including the introduction of internal peer-review and surveillance (48). P4P has driven major changes in the roles and organization of English primary care teams (50). The structural specialization due to the P4P gave a better organizational competence (49), but may also have given non-incentive activities and patients' concern less clinical attention (50). Campbell et al (51) have interviewed 21 family doctors and 20 nurses in 2007 and found that P4P has changed the dynamic between doctors and nurses and the nature of the practitioner- patient consultation.

3.4.11 Doctors autonomy

Have the financial systems impact on the doctors autonomy and internal motivation in primary care? And if so are the doctors conscious these impacts ? Two ethnographic studies have been carried out (52,53) and they identified a real shift towards the delivery of a more biomedical, disease-orientated model of care occurring in response to the imperatives embodied in the new general medicine services contract (53). Checkland et al (53) performed an ethnographic observational study in four practices over a period of about 6 months. They comment that GPs' traditional self-identification as 'holistic' or patient-centered practitioners, has been challenged. As occurred in 19th Century hospitals, patients in modern general practice are increasingly being identified by disease labels, and their care is being determined

by those labels. In spite of the real changes occurring in their practices towards a more biomedical model of care, the doctors in the study seemed unaware of this, denying that significant change had taken place and locating any change at the margins of practice (53). McDonald et al (52) also did an ethnographic study, lasting four months. They included 12 GPs, nine nurses, four healthcare assistants, and four administrative staff. Nurses expressed more concern than doctors about changes to their clinical practice, but also appreciated being given responsibility for delivering on targets in particular disease areas. Most doctors did not question the quality targets that existed at the time, or the implications of the targets for their own clinical autonomy. Implementation of financial incentives for quality of care did not seem to have damaged the internal motivation of the general practitioners studied, although more concern was expressed by nurses (52).

3.4.12 What are doctor`s experience with P4P on their workday?

The implementation of QOF by P4P was intended to improve quality, reduce disparities in care and enhance access, but also to improve the working lives for GPs. By questionnaires filled out from 2179 GPs in 2004 and 1378 in 2005 Whalley et al (5) and Gemmel et al (54) have looked into the GPs evaluation of working life before and after 2004. They found that general practices may have responded to the 2004 contract by increasing staffing levels, with nursing staff absorbing a higher proportion of the clinical workload and doctors focusing more attention on chronic and preventive care (54). Despite negative consequences for workload and autonomy, the job satisfaction seemed to increase after introduction of the new contract (5). GPs reported working fewer hours with a higher income, and their expectations regarding the impact of the contract on quality of care had been exceeded (5). McDonald et al (55) used an in- depth qualitative case study in two general practices in England to investigate mechanisms and perceptions of control following the implementation of QOF in general practice. Although the sample was small, they found interestingly and in consistence with the literature on motivation and surveillance (40), that tensions were greatest in the practice with more top-down approach. They found that attitudes towards the contract were largely positive, but discontent was higher in the practice which employed a more intensive surveillance regime and greater amongst nurses than doctors (55). The last may be due to GP having more clinical autonomy than nurses.

3.5 Different structural devices of QOF

3.5.1 Exclusion of patient from the QOF

The literature points out the importance of having the possibility to exclude patients from quality measures in the QOF program, to protect against inappropriate - or over treatment of patients (9,10). It is stressed that this together with a possibility to reach full score without having every target reached, is doing the system more defensible, robust and less intrusive in the patient doctor relation (9,10). Although it also gives the doctor the opportunity to cream-skimming by excluding the patients being non-compliant (9). Dalton et al (56) used the method of three cross-sectional analyses with data from EMR of all patients with diabetes in 23 general practices in Brent in 2004/2005 and 2006/2007. They found that the patients excluded from the pay-for-performance program, may be less likely to achieve treatment goals and disproportionally came from disadvantaged groups. It concluded permitting physicians to exclude patients from P4P program may worsen health disparities (56). Doran et al (9) extracted data from 8229 English family practices in 2008-9 and Doran et al (10) analyzed data extracted from 8105 family practices in England in 2005-2006. These two bigger studies (9,10) concluded that relatively few patients were excluded for inform dissent, suggesting that the incentives activities were broadly acceptable to patients (9), and that exception reporting brings substantial benefits to P4P programs, providing that the process is used appropriately. In England, rates of exception reporting have generally been low, median of 2,7%, (9) with little evidence of widespread gaming (10).

3.5.2 Piloting the Quality indicators

The QOF, a P4P- scheme, as said earlier, was introduced in 2004 by the government in the United Kingdom. It consisted of clinical and organizational quality indicators. The original QOF indicators, and all subsequent indicators, were introduced without piloting (57,58,59). In 2009 a new way of developing clinical indicators for QOF was introduced, led by the National institute for health and Clinical Excellence (NICE) (57). The importance and value of piloting are emphasized in terms of an opportunity to identify unintended consequences of potential QOF indicators in the «real world» settings, with staff who deliver day-to-day care to patients (58). Four particular types of unintended consequences had been identified: measure fixation, tunnel vision, misinterpretation and potential gaming. ‘Measure fixation,’ an inappropriate attention on isolated aspects of care, appeared to be the key unintended consequence. Hannon et al (59) interviewed 57 staff members in 24 practices after a single

palliative indicator was piloted in 6 months in 2009. The aim of the study was to gain the views and experience of general practice staff on whether the inclusion of a single incentivized indicator to record the preferred place receive end-of-life care, would improve the quality of palliative care. They conclude in their study by saying that the most appropriate time to ask a patient about end-of-life care is subjective and patient specific, and therefore not lend itself to an inflexible single indicator. Focusing on one isolated question simplifies and distracts from a multi-faceted and complex issue and may lead to patient harm (59). In particular, if the palliative care indicator had been introduced without piloting, this might have incentivized poorer care in a minority of practices with potential harm to vulnerable patients (59). The UK government currently spends over 1 billion pounds each year on QOF. Each UK pilot cost £150,000 (0.0005% of the overall cost). The act of piloting indicators is therefore value for money as it identifies implementation issues of acceptability and unintended consequences, as well as technical reliability and feasibility that can be addressed and rectified prior to national roll out. Moreover, it highlights indicators that should not be included. An indicator testing protocol must act as a foundation stone for the field, testing and development of country specific quality indicators for pay-for-performance or quality improvement schemes. Local adaptations of this protocol could be used by policy-makers and researchers to empirically test the likely effect of implementing indicators. Whilst national in scope, the lessons are, we hope, therefore generalizable for an international audience (57).

3.6 Ethical aspects of the QOF

I have sought in PubMed, Best practice, Up to date and BMC finding only a few articles from the UK about ethics and P4P. I have found some few from UK and some articles from America and one written by authors from New Zealand and UK. I find this very interesting. I'm sure there are more articles from UK covering this issue, there must be, but they are apparently difficult to find. Why are there so few reflections from the UK on ethics issues linked to P4P/QOF/quality improvement? Tapp et al (60) reflect on enabling general practitioner in UK to judge ethical dilemmas under the quality improvement framework. The reflection is triggered by the ethical dilemmas, quality improvement projects that push boundaries, gives. They had no clear conclusion and saw the difficulties in establishing a system of ethical approval in increased workload. More bureaucracy will create barriers to implementing quality improvement, at the same time as they saw that there is often no mechanism in place within Family practice to deal with these issues. They conclude it is time to encourage more ethical reflection in practice. GPs should review each project in terms of

their intentions, the impact on the patient and the data collection methods to determine if any aspects of the project need further consideration. This reflection should be guided by medical ethical principles (autonomy, utility and justice), recognising that individual patients' views on treatment and autonomy may be in conflict with quality improvement (QI) schemes. A flexible approach must be adopted, depending on the nature of the practice. Larger practices may want to establish a more formal internal process to review their QI projects and approve them as a practice. In the future, patients' perspectives on their treatment and involvement in decision-making, etc. can be incorporated into QI schemes, in order to ensure that patients' views on utility and autonomy are not violated (60). Buetow and Entwistle (61) conclude in their reflections about pay-for-virtue as an option to improve pay for performance, that concerns about current P4P schemes erodes professional values and undermine moral motivation might be ameliorated to some extent by 1) investment in pre-qualification moral training and post-qualification support for ongoing moral reflection and development.2) Modification to current pay-for-performance schemes to improve the assessment of individual patient care and/or 3) the addition of indicators of virtuous practise to assessment frameworks (61). In the article from the US, Whraham et al(62) conclude that current arrangements (P4P) are based on fundamentally acceptable ethical principles, but are guided by an incomplete understanding of health-care quality. Furthermore, their implementation without evidence of safety and efficacy, is ethically precarious because of potential risks to stakeholders, especially vulnerable patients. They conclude by proposing four major strategies to transition from risky pay-for-performance systems to ethical performance-based physician compensation and high quality care. These include implementing safeguards within current pay-for-performance systems, reaching consensus regarding the obligations of key stakeholders in improving health-care quality, developing valid and comprehensive measures of health-care quality, and utilising a cautious evaluative approach in creating the next generation of compensation systems that reward genuine quality (62).

Snyder et al (63) says in 2007 in their ethics manifesto that by framing the discussion of P4P in terms of ethics and professionalism, both in the context of the individual patient and for society, they hope to move the debate forward with a patient centered focus - one that puts the needs and interests of the patient first—as P4P programs evolve (63). They continues that P4P programs may have the potential to increase overall quality of care, when aligned with the ethical obligations of the physician to deliver the best quality care to her or his patient, but the risks aren't that incentives could result in de-selection of patients, “playing to the measures”, rather than focusing on the patient as a whole, misalignment of perceptions

between physicians and patients, and increased unnecessary care and costs, have the potential to harm access to care, continuity of care, patient-physician relationships, and care for those patients with complex chronic conditions (63). Nelson (64) address the ethical challenge in making an incentive system rewarding quality and efficiency, both being subjective terms. He asks what is quality and for whom? And how can the doctor manage to balance the demand on effective deliver of care and at the same time serve the patient with the medical ethical principles of doing good and no harm. He wants more consciousness about the possible unintended ethical consequences of an incentive system.

3.7 Does the QOF work – has it given the desired results?

Campbell et al (38) conducted an interrupted time-series analysis of the quality of care in 42 representative family practices, with data collected at two time points before implementation of the scheme (1998 and 2003) and at two time points after implementation (2005 and 2007). At each time point, data on the care of patients with asthma, diabetes, or coronary heart disease were extracted from medical records; data on patients' perceptions of access to care, continuity of care, and interpersonal aspects of care, were collected from questionnaires. The analysis included aspects of care that were and those that were not associated with incentives (38).

Campbell et al finds that between 2003 and 2005, the rate of improvement in the quality of care increased for asthma and diabetes ($p < 0.001$) but not for heart disease. By 2007, the rate of improvement had slowed for all three conditions ($p < 0.001$), and the quality of those aspects of care that were not associated with an incentive had declined for patients with asthma or heart disease. As compared with the period before the pay-for-performance scheme was introduced, the improvement rate after 2005 was unchanged for asthma or diabetes and was reduced for heart disease ($p = 0.02$). No significant changes were seen in patients' reports on access to care or on interpersonal aspects of care. The level of the continuity of care, which had been constant, showed a reduction immediately after the introduction of the pay-for-performance scheme ($p < 0.001$) and then continued at that reduced level. Campbell et al conclude that between 1998 and 2007, there were significant improvements in measurable aspects of clinical performance, with respect to the care provided for three major chronic diseases. The initial acceleration in the underlying rate of quality improvement after the introduction of pay for performance was not sustained. If the aim of pay for performance is to give providers incentives to attain targets, the scheme achieved that aim. There may have been unintended consequences, including reductions in the quality of some aspects of care not

linked to incentives and in the continuity of care (38).

Steel et al (65) carried through a national structured survey questionnaire with face to face interviews. They included 8688 participants over 50 years old or more. They found a large gap between recommended care and care that is actually recieved. They concludes that routinely availability of information on performance, used to actively monitor performance, is associated with better care (65). In England, performance monitoring through the general practice pay for performance contract has been linked with improved care for included conditions (65). Making information on performance available for a wider range of conditions is an essential component of quality improvement. Chronic conditions that affect quality of life of older people must be included in future revisions of the contract (65).

Sutton et al (66) estimated the intended and unintended consequences of the Quality and Outcomes Framework (QOF) using dynamic panel probit⁸ models estimated on individual patient records from 315 general practices over the period 2000/1–2005/6. Unintended positive consequences may be investments in quality that have benefits outside the focus of the scheme and negative consequences may include effort diversion and gaiming (66).

They focused on annual rates of recording of blood pressure, smoking status, cholesterol, body mass index(BMI) and alcohol consumption. The BMI and alcohol consumption were un incentivized risk factors. They used these risk factors as a comparison group for the incentivized indicators and estimated the spillovers to these unincentivized activities for the targeted patients.

Most studies focus on incentivized activities only. Sutton et al (66) criticize other studies for not looking into unincentivizes activities, not doing formal analysis or being too small. They criticize to of the articles used as references in this these. I will therefore refer their criticism after a short summary of the two articles;

Campbell et al. (34) examined the quality of care for three of the diseases targeted by the QOF (CHD, asthma and diabetes). They compared the observed trends in incentivized and unincentivized indicators for these patients with projected trends based on measurements taken in 1998 and 2003. The 2003–2005 increases were significantly higher than the projected trends for asthma and diabetes but only marginally so for CHD. The rate of improvement for the incentivized indicators did not differ significantly from the rate of improvement for the unincentivized indicators. The study by Steel et al. (4) is the only one to consider

⁸ In statistics, a probit model is a type of regression where the dependent variable can only take two values, for example married or not married

unincentivised and incentivized indicators for targeted diseases (hypertension and asthma) and different quality indicators for two untargeted diseases (osteoarthritis and depression). There were significant increases on the incentivized and unincentivised indicators for the targeted diseases between 2003 and 2005. The indicators for the untargeted diseases did not change significantly. Sutton et al's (66) criticism says that Steel et al. (4) did not undertake a formal difference-in-differences analysis and Campbell et al.'s (34) analysis relies on projected trajectories based on only two time points and with no allowance for projection uncertainty. Both involve small numbers of practices (n518 and n542, respectively). Sutton et al (66) goes on and claims that none of the previous studies has exploited fully the presence of disease and indicator combinations that are or are not incentivized.

They also claims that no previous study has compared the effects with the costs of the scheme and their estimation are that the positive spillovers contributes to reduce the costs by a factor of two (from an estimate of 50 pounds per record to 25 pound per record).

They found that the incentives improves targeted performance and has positive spillovers for targeted patients. They also found that the responses of providers depend on the design of the incentives. Finally, as said showed how the existence of positive spillovers affects estimates of the unit costs of the scheme (66).

3.8 Does P4P give value for money?

Walker S et al presents a cost-effectiveness analysis of nine QOF indicators with a direct therapeutic impact with data from all English primary care practices (67). They used an economic analysis and found that for most indicators that can be assessed, QOF incentive payments are likely to be a cost-effective use of resources for a high proportion of primary care practices, even if the QOF achieves only modest improvements in care. However, only a small subset of the indicators has been considered, and no account has been taken of the costs of administering the QOF scheme (67). There is no evaluation of the cost for reduced focus on not measured conditions. Sutton et al (66) found that the incentives improves targeted performance and has positive spillovers for targeted patients. They also found that the responses of providers depend on the design of the incentives.

4. Results/Analysis of the interviews

The main findings from the two interviews collated along two axes as I perceived it; positive - and negative effects of the Quality and Outcome Framework. I have structured the main findings under these two categories, and sorted further into the different aspects of the positive - and negative effects respectively: systematics as the one sub headline in the first category and Doctors versus patient centred, overtreatment, political tool and others in the second category. As I was very interested in the participants opinion about QOF's effects on Quality and QOF and ethical issues, I have kept the reflection on these two themes separate. I start the analysis by presenting what transpired on these two themes.

In the first interview the participant is presented as GP and interviewer, me, as I.

In the second interview the four participants are presented by different letters; L, A, AA and E while I as the interviewer by KJL.

The interviews – with the experienced GP and the group of younger GPs addressed many similar topics and challenges, sorted under the described categories. What I found interesting was that the experienced GP reflected more on QOF as a system, while the group of younger GPs mostly reflected in to the systems details – from the inside of the system and in the frame of the system. May be not strange, as the experienced GP knew “before and after” the introduction of QOF, but interesting that it was difficult to get reflection about the system as a whole from the younger doctors and as they said they thought this maybe was about a generation shift:

L: *“I think it has something to do with the generation as well. Because I think we are coming from a generation that has been trained to be much more structured or more aware of protocols – more accountable there is been a shift.”*

.....

L: *“I don` t think it bother us so much because it`s just the way we work and it is the way that we always been working. “*

4.1 Quality

How is the quality of family medicine and the quality in the patient meeting affected by the introduction and the use of QOF? These are huge questions, difficult questions and for sure there are no definite answers to it. But the reflections from GPs working with the framework were interesting.

In the first interview the GP said that QOF can be supportive to ensure quality by securing structure to important medical measurements. Some of the statements and reflections about quality could also be sorted under positive effects as I have done in chapter 5.

When he was to define quality the answer was:

GP: "...Quality is about doing my personal best for patients, burying in mind the patients preserved needs and my preserved concept of what's best for the patient, hypotension and diabetes being example. It is about being rigorous, it is about being professional, it is about treating the patients across the board whatever I feel about them myself, leaving my personal feelings asides or rather using my personal feelings as a tool within the consultation, rather than at something that alienates me from the patient. It's about always trying to improve, so it's about learning with my colleagues, it's about learning from research, it's about keeping my eyes out and my ears out for how I can improve health care, it's about questioning my own practise, those of my colleges and it's also questioning the whole system. The system we have whether it is in the practise or wider in the health service and questioning how that impact on quality. It's many headed..."

During our conversation I understood that he meant that to improve, you have to be aware of giving the patients attention and respect, and you have to be able to compare measurable results. To be able to compare, you must have something measurable. QOF gives measures in many aspects, it gives the possibility to measure changes and to give you results telling you whether you have or haven't done a good job, compared to yourself, compared to your colleges or even wider bench marketing. The side effects are that all the not measurable issues and topics can be undermined and threatened. It makes it difficult and even more demanding for the GP to be aware and manage quality and quality improvement.

The group in the second interview addressed quality in several ways. They agreed on that the term quality is depending on the definition:

A: *"The - the question to me is who define quality? Is it quality of output from the point of view of better service for the individual patient? Is it quality of outcome in terms of a government turning on saying "well in our government few people have the cholesterol of 5 or more..."*

Three different perspectives of Quality where addressed first they talked about the perspective of the employer and the government their focus on "get what they want". They talked about the way the doctor's recognise og look at quality and they addressed the way to establish and making visible the appropriate level of service the patient should expect.

When we discussed further quality and QOF one of the GPs in the second interview gave an example on how she had experienced QOF to catalyse quality improvement:

L: *" one of the practise I have worked as a salary doctor, was a practise that previously had been a single handed doctor who had had it for many many years and had a very bad reputation for being a bad practise. It kept its patient because it was a very deprived area and they just stayed with the practise because they knew the doctor. The two partners who took over, took me on as a salary doctor and you know they were filling in with QOF at the time because of the history and actually having QOF really catalyst them in to action – to get their diabetics a bit more under control, they got their hypertension under control they had not been controlled for a very long time aam...and I think that is quite a good example ."*

The literature gives many definitions to quality, some of them presented in chapter 1, Background. The participants here gave a description and definition to quality from their clinical view. That's important and should be given much thoughts back into the theoretical definitions - do the definitions cover what is the main issues in the everyday delivery in a surgery for UK - or for that matter the rest of the world GPs?

My impression was that the participants gave QOF credit for giving a frame and "security system" for ensuring a certain level of quality, but is this to a level of justifiability or does it secure a more ambitious level of quality improvement? After analysing the interviews I'm not sure it improves above the level of justifiability. They told me about a not perfect system and they also told about a system which needed a constant development and improvement, to reduce side effects and the perverse effects. But after all they gave me the impression of a fairly good system for quality assurance, securing a minimum level. They also asked the question whether it rather is about control and a political way of showing that the government

take action, cf also 4.4.3 political tool, than really about quality improvement. In the interviews the participants gave after all an overall fairly good evaluation to QOF – as a system lifting every GP up to the same bar.

E: “... I have no doubt that this process has helped tightening up and then recording of a lot of conditions and the management for a lot of conditions.”

But they are also unison about the QOF`s limitations. The quality lift is to a certain level, and for the practices already performing well, it is a short way to experience the QOF as a hinder to further quality improvement:

E: “They are obligated to various practises who perhaps are delivering quality service anyway are now just ticking boxses. They are delivering rather than improved the level of the quality of the practice, that`s what it feels like.”

Or:

E: “It seems that QOF helps to a certain degree and then it starts interfering”

And the side effects – or may be it is need for an even stronger word – the perverse effects were among other described as:

E: “I think what it`s feels like it`s doing its trying to put – it - we are ticking boxes on forms for QOF but it feels like it tries to put patients into boxses as well. You know ... which just doesn`t work is -- one individual patient which comes in with depression, is going to be completely different from the next ones who comes in with it. To make them try to fit in to this protocol you know -- this is what`s we must do but it really never quite work”

4.2 Ethics - Ethical issues and the QOF

Can ethical dilemmas be encountered due to QOF?

As medicine is a value-based practise, Ethics is, and should be, an important issue for every doctor. I wanted to know if the GPs encountered any ethical dilemmas due to QOF.

During the interviews the GPs touched ethical aspects through describing different sides of their practise and different ways they saw QOF interfering with them being a “good doctor” and the way they wanted to perform and had ambitions for their job. They pointed to several

ethical issues related to the QOF, such as feeling obligated to introduce questions or schemes into the consultation that never would have been brought in, if it wasn't for QOF. This was described as interrupting the consultation and even in danger of interrupting the relation between patient and doctor. This was addressed by one of the GP as an danger of reducing the doctor's empathy. In both interviews they addressed the problem of introducing themes that the patient did not ask them for, and as addressed in the first interview ; *"... to introduce a problem to the patient that he or she did not see as a problem, with the risk of introducing new problems into a patient's life."*

This would also introduce the danger of pushing the patients agenda aside and give the QOF/doctors agenda to be the main issue of the consultation. They described feeling pressured to "shoehorn" in QOF targets. Pressured by the authorities system, but also pressured from colleges especially the GPs working as salaried doctors. They felt like having an expectation from their employers (GP partners) on delivering the QOF targets, and less free to decide when QOF targets where not adequate to use or bring into the consultation.

As noted earlier, there are some differences in the way the ethical dilemmas are addressed in the two interviews. In the first interview the ethical dilemma was addressed due to in what way the QOF intrude into the patient's life, putting the patients need away and forcing the doctor to give advises or give treatment due to the measurements and demands in the QOF, rather than a holistic medical approach. The younger doctors addressed the ethical dilemmas also due to driving the focus away from patient centred care, but they went more into details of the system and the way they managed to handle the different demands – whether it was ticking the boxes or filling out schemes into a consultation on a false or honest base. They often felt that the QOF schemes or targets really did not fit in, so to reduce the disturbance they found themselves introducing schemes or targets with falsity as something the patient was in need of.

In the second interview they also problematized the huge amount of recourses going into the framework, financial and human.

Another theme touched in the first interview, but more thoroughly addressed in the second: The link between QOF score/ticking boxses and the money connected. In the second interview they asked the question - maybe it would be easier to defend the QOF demands if it were not connected with money? Is it possible to connect targets to other incentives? All the GPs in the second interview were salaried or freelance doctors.

When asking them directly if they encountered any ethical dilemmas due to QOF they pinpointed differently and not only in a negative way – in the second interview it was pointed out from one of the GPs that it also made here feel like a better doctor in an ethical way:

AA: *“ I think there are times where I feel I’m being a better doc. revenge of QOF because it helps me give me the prompts and helps me to base on guides and it helps me to give best practise and I think there is connection make you feel that you are more of an ethical doctor because you are being a better doctor in that sense...”*

But as the GP continues she got into more negative issues about QOF and ethics:

“... I think there are other times where I... it - - I think it can do patient a disservice as you mentioned before - the PHQ9 the depression screen I think it’s a prima ex. And then if I have to be honest the QOF requirements I would not being a good doctor to that patient in front of me who’re distress and doesn’t want get the questionnaire showed in her face and if I do give them the questionnaire I find myself present it in a way that I am not honest with the patient that this is a QOF requirement... I actually I don’t feel honesty to need I feel like leaning on the patient from the chat we have had and that’s disgusting.”

E: *“ But it’s not fitting into you practise...”*

AA: *“ No and I am kind of introducing something with a falsity and I actually have to lie which stresses it up and I that makes me feel quite uncomfortable and that’s just one word a very simple example.”*

.....

In the first interview the GP answered the question about ethical dilemmas due to QOF in this way :

GP: *“... I think there is an ethical dilemma when my agenda clashes with the patients agenda very simple actually on that level ahh as to what I am gone spend my time with in the consultation I am going to – how I am gone lead the consultation or to what extent I am gone let the patient come up with their agenda that’s an ethical dilemma, it’s a very fundamental one actually ahh and I think where I am thinking about my money rather than what the patient wants or even is best for the patient is an ethical dilemma...”*

The theme about being pressured by the QOF targets and feeling obligated to push them into the consultation, even though it disrupted the consultation, and sometimes bring in quite another focus:

E: *“There is a danger when you are pressured from whatever source -- that your empathy is reduced so that’s an extreme example asking do you have any allergies when someone are coming up with a comment like that...”* (my comment: earlier in the interview there were given an example where a student interviewing a patient got a comment from the patient “that’s the worst thing that ever happened to me, it’s like being raped” and the medical student, only stressed by getting through the QOF targets says: ” and do you have any allergies ?”)

The GP continues:

E: *“...But I think quite possibly that at one point we have all done it that we have tried to shoehorn in QOF target at the end of the consultation purely because we feel pressured to do it, even then we know that it’s not right but equally we are pressured so it must come across very badly.”*

In the second interview they addressed and problematized several times the financial costs and the huge amount of resources used to administrate the incentive system and addressed this as an ethical issue. As I perceived what they said, they felt that all the huge amount of money and resources going into administration and policing QOF were to be questioned whether giving enough quality improvement to be defended. Into the administration they described controllers in every practise (a college), managers of the IT systems, people in NHS,NICE and other organisations administrating the QOF system.

L: *“The question about ethical is about whether the money that QOF actually takes to being implemented is actually being met and honored with improvement in practise.....are the money sort of actually being reflected in improvement practise ? I simply have no answer for that...”*

4.3 Positive effects of P4P

4.3.1 A more systematic approach

The perceived positive effect presented and pinpointed by all the GPs was that the QOF had provided them more systematics into their everyday work and lifted every GP up to the same bar. They also talked about the experience of having more evidence based - and more secure

treatment, for the patients on the targeted medical conditions. To their experience QOF really helped them to get attention to many important issues and in a systematic manner, by screening everyone and not just some patients. They also agreed on, given the frame of big workloads and a high tempo in the everyday, it is a good help and security having a system reminding them of important measurements, and doing so quite systematic. They problematized the control aspect of the measurements but concluded that the positive effects exceeded the negative as long as it is based on medical evidence with hard endpoints. Everyone gave me the impression of being worried that QOFs positive effects only extends to a certain level and when doctors reaches that level, all of them mentioned concerns about the negative effects the system can provide cf. 4.4 negative effects.

The QOF is a system that apparently helps the GPs to systematic follow up their patients in the matter of reminding and regulating them to control (important) defined measurable parameters. A very important success factor for the system, according the GPs, was a good and advanced IT systems. They present it as an absolute prerequisite for the framework to work and an absolute premise for further developing of the framework.

When addressing the evidence based and hard end point driven measurements in QOF as an important reason for why they described the system in positive manner, was expressed by one of the GPs in this way:

GP: *"...But that is an example where process is very much linked to outcome and where there is research evidence that the controlling blood pressure has an effect on outcomes in terms of cardiovascular morbidity and mortality."*

So as long as the measurements are linked to hard endpoint it can easily be defended, they all agreed on that. The GP continued:

GP: *" I have no problem with that actually. And I find it very difficult to argue against, I think it is impossible to argue against and so I am happy with that aspect of QOF."*

The systematics was elaborated by one of the GPs

GP: *"...There were people who just fell out of the system for whatever reason and people do succeed in "reggeling" through the gaps and the... and the net and quite successfully sometimes and I was horrified that we had patients who haven't had their kidney function checked who were on Lithium for instance and there were three of them and I looked at them*

horrified and I went into their notes and one of them I knew the other two I didn't. And they had obviously tried to very hard to "reggel" out of and we let them."

The way QOF helps the GPs remembering and securing the delivery of good practise was address several times by different GPs in the second interview:

L: "...You start, if you start working as – even if you salaried even if you're not a partner, but you got your own patients you can start seeing the relevance of it..." ... but you can't remember everything and having someone sitting in the background saying: actually you saw this patient with this particular problem and you forgot to do this or it wasn't done for whatever reason- how about you going back and review - what – that's useful."

In the second interview they also had a discussion about whether the QOF had to be connected to payment. They saw the negative effects up to perverse effects as long as the system gave you money for ticking boxes and filling out schemes. At the same time:

AA: "... The issue with that starts to be- there if you are being paid to see a patient, not to actually do anything useful with the patient then that is where the problem is. So the only way to address that perhaps is to start bringing in some targets of some kind."

4.4 Negative effects of P4P:

As the positive effects of QOF were addressed as systematics by all of the GPs, I also found many of the themes I have sorted under negative effects expressed by everyone. As the GP participating in the first interview also had worked before the introduction of QOF, he gave reflections about the changes introduced with the framework. Changes that had affected the interaction between the patient and the doctor, as he presented it. He problematized changes in focus in the consultations, where the GP is dragged towards the computer and ticking boxses, away from the patient. Also the GPs in the second interview addressed this issue about how QOF is putting the service system (QOFs boxses, schemes and guides) in the centre and pushing the patient agenda more or less aside. I have sorted the negative effects into subgroups, and this issue is described in 4.4.1 under the headline of doctor versus patient centred care. On the question about whether QOF has affected the family medicine as a professional science, the GPs in the second interview gave an interesting answer to this question with a both yes and not exactly no, but more like "it depends on the GP" answer. As I understood them, they meant that if you let yourself as a GP be very obligated by the

protocols (in QOF) and not letting your own knowledge and experience be the foundation for your practise – yes, then it has made a change on the family medicine as a science. All of the GPs stresses that a rigid system used on a service and a science, where the target audience and the science core are of huge variation, and naturally must be, can be pretty problematic. So the way the doctor use the system is essential. The variations are due to differences in gender, age, social situation, economic and educational situation, genetics, predisposing and so on. So as the QOF can be a tool for systematics (cf 4.3) it can also be a tool in risk of deindividualising the medical treatment. As one of the GP said in the second interview:

“One size fits all doesn’t work “

All of GPs addressed the problem with overtreatment. Just as the target goals in the QOF not necessarily fits all, it can even be quite wrong to meet medically for some patients. In the first interview the GP was worried about sometimes the drive for reaching targets could give the patients reduced life quality, everyone was worried and mentioned that even worse they could ending up killing the patient. In both interviews they addressed the problem with being pressured to treat the targets and reach the targets both from colleges and the system.

In both interviews they also addressed that QOF is a potent political tool. They talked about how politics are creeping into both the recommendations and the targets – with the risk of putting the medical science and evidence based medicine aside.

In the interviews there were also commented and reflected on other themes like transaction costs in relation to cost effectiveness, gaming the system by fixing or holding back results not achieving the set targets of QOF, but with less conclusiveness in the discussions.

One theme mentioned by one of the doctors was her experience of mounting pressure from the authorities resulting in a constant raising working load:

E : “... *I am feeling the pressure just as a salaried doctor so I can just only imagine what the partners feels. hmm and I mean QOF is just one part of that, **the pressure on us now is forever mounting,...***”

All the participants agreed on this statement.

4.4.1 Doctors versus patient centred care

In the first interview I asked the question: Do you think QOF has influenced your way of being a doctor, or your practise in any way, professional way? I was especially interested in what his thoughts was about this since he had worked many years before QOF was introduced, so I will quote some of his answers telling about his experience. As I heard him he described a change in the interaction between patient and the doctor:

GP: *“Yes it’s made – I think I’m going to go back to what I said initially it’s made me more systematic, but it’s also inevitable diverts me and I used that example of the patient who I ended up being more concerned about how filling out the questionnaire, than about her depression and that is inevitable to that it diverts. I think that it just in purely in terms of the consultation, it actually diverts me to the computer screen rather than to the patient and I had to get into the habit of pushing myself away from my desk and looking directly at the patient if I am not actually on the screen...”*

...”...And that’s another conflict between patients centred care which we like to pride ourselves on. And the realities of the agenda that has been set for us, which is actually ehh.. some people will argue at the patients centred in terms of outcomes. But during the consultation it is very very difficult to not have what is essentially the doctors agenda intrude up on what the patients may have come with. There is always going to be a conflict there of course, but it makes it’s starker, I think. “

.....

And when we talked further about the changes QOF introduced he underlined again the change in focus and the way he experienced to be drawn or forced into ticking the boxes:

...GP: *” ... Because one of the HUGE drawbacks of QOF is that you end up treating the boxes you tick the boxes and not the patients and that brings me back to the intrusion into the consultation and it’s like so many things in health care and certainly in the health service in Britain there are series of conflicts, of dilemmas conflicting, pulls and pushes and that is another example of conflict which is not resolvable, but it needs to be recognized and named and managed.”*

In the second interview they also addressed the problem with QOF giving a more service centred (“Box centred”) care than patient centred care, and in both interviews the scheme for depression score mentioned and problematized:

E:” ... A really good example of that is I find is the questionnaire we are asked to do for the depression. And you see you got a patient in that you stop to discuss issues and that will automatically move into a very complex discussion usually which will last slightly longer than you expect a normal consultation to last. Somewhere in the middle of that you are expected to say: could you fill out the score please, so that I can quantify how depressed you are - it never, it never sits well in the consultation at any point - ever. And sometimes you feel actually have to do this for QOF if don't do this I ...— but I can't fit it I can't shoe horn it into this consultation. So quite often I end up to having to separate it out of the consultation. I say; hey look as a part of the way we gage how well you are at the moment I would like you to fill it out and bring it back to me... “

.....

When I asked whether QOF has done something with the family medicine as a professional science they gave this yes and no answer exemplified by this quote:

A: “There is a perception that if you following a protocol and if you consider that /it as more than a guideline if you feel a very strong obligation to follow protocol that the computer throws at you, rather than use a combination of your knowledge and your experience there - certainly the sense that I have got from colleges- is that they feel that it detracts from their personal sense of professionalism.”

Everyone gave approval to this statement and it was even more underlined and stressed:

E :” I think what it feels like it`s doing it`s trying to put – it - we are ticking boxes on forms for QOF but it feels like it tries to put patients into boxses as well ...”

E: “**One size fits all doesn't work** “

The second unison challenge addressed by the participants was QOF pushing the doctor`s to over treat their patients:

4.4.2 Overtreatment:

In both interviews the problem with the systems expectation about reaching the targets in QOF was underlined, but also the pressure from the colleges about reaching targets, could make them do more than what was strictly medical necessary, or medical recommendable , even sometimes at the patients expense.

E: *"... the communist one is to thrown out is the HbA1c guides or the QOF targets for 7 % or I don't know if that's changed slightly now but the used to be right everybody - every diabetic we should be aiming for a HbA1c 7% or for 48 or less, and that is your target. Well that's not always appropriate if you have got an elderly diabetic and you force their HbA1c down with whatever therapy you need to do to get that you might start to give them hypoglycemia."*

A: *"You might kill them"*

In the first interview overtreatment was addressed like this:

GP: *"No, there are of course downsides of that, blood pressure can be over treated and produces its own - its own morbidity even mortality. I'm particularly interested in elderly people and over treating blood pressure is common around probably more common in older people than elsewhere in the population and of course you get false and things like that."*

I: *"Do you correlate the overtreatment to QOF?"*

GP: *"Not entirely no, I'm saying it's a danger"*

.....

When we continued the conversation, the GP gave two examples – overtreatment to diabetic patients giving hypoglycemia and treating hypertension too heavily, which ends up giving the patients postural hypotension. He stresses that the knowledge about this is too little and it have to be given attention:

I: *"So they're life quality is reduced?"*

GP: *"Yes, and someone really needs to research that."*

I: *"But to get the QOF then satisfied or at the best level you have to keep on regulating them as hard as this?"*

GP: *"Yes that's right, you do..."*

In the second interview they even felt the pressure as salaried doctors in the way of being more obligated to deliver on behalf of the partners.

L: *"... I have certainly felt ... pure pressure from my colleges, even though I for me, you know, beyond the benefit of the patient there is no financial incentive for me to undertake the QOF"*

but there is a huge amount from the colleges in the practise and from the manager say that you served this patient but it did not have a QOF target – especially when they are close to April – that`s the time every year the QOF is completed. I don t have the guts to say : oh well I have not hat was not important if there was an exceptional consultation that the persons is very depressed or something and you know - it is really not proper you can justify it, but otherwise I think you have to be a more senior GP to be able to stand up against ...”

E: *“They say this is a part of that job and you must do it. So you feel that your autonomy is slightly removed as a salaried doctor.”*

The pressure from colleges is also mentioned in the first interview:

GP: *“But then in the partners meeting someone says you know _ the following things haven`t been done they were last seen by doctor B you know so it`s a little bit of name and shame. And there is a lot of peer pressure and that is fine it`s how we work as a team: what I do and don`t do effects everybody else`s everybody else`s well it`s the only way to do it but there is a lot of pressure on us to tick the boxes.”*

4.4.3 Political tool

The QOF system is also described by the participants as also being a political tool – and a potent one. All of the GPs talked about how politics are creeping into both the recommendations and the targets – with the risk of putting the medical science and evidence based medicine aside. Having the possibility to be lobbying on important medical issues where stressed as a positive democratic possibility, but then again problematised as a potent way of getting capitalists or even quite other negative or suspect intentions into the medical science. The QOF systems indicators and guides are based on evidence and hard facts in medicine, but at the same time every doctor knows that we only have evidence for parts of our science, therefore much is based on experienced science and not at least knowledge about - and deep respect for humans. What kind of forces do we have to be aware of, in a system giving the authorities the possibility to interact with the health service delivery directly, directly into the consultation?

In the first interview the GP stated that the development of QOF indicators and measurements are now getting more and more into a political issue as the measurements are changed around or they are trying to put more into the measurements:

I: “... And you said that if you get more measurements into the QOF it’s a risk as well-“

GP: *“This is where it becomes a political issue, now you will hear people deny this - it is clearly very political I’m not saying it is only political but there is a very strong political element in it and if the Norwegian government is thinking of setting this up you got to watch out for this – because everyone wants to jump in on the band wagon.”*

In both interviews the third month injection as recommended contraceptive, was mentioned and criticised. All participants questioned the choice of this recommendation and hinted it could be more about economy and politics, than good medical advice.

From the first interview:

I: “So the government will prefer that third month injections ...”

GP: *“Someone thought this is a good idea and pushed it and I am sure they pushed it very hard. And pushed it and pushed it and then it has got into QOF.”*

The same issue with the contraceptive advises was also addressed in the second interview:

L: *“Well actually that’s an interesting point because this was based on the NICE guidelines and when I said actually the best type of contraception and long active regressive contraceptive which is a depot injection or a coil which is –ok there’s less likely that you forget to take it, but because it is in there- but it’s actually much more invasive than taking a contraceptive pill and I am surprised about NICE guidelines and be then the QOF said actually we have to convince everyone to be put on them because that’s very much cost effectiveness rather than perhaps what the patient want or may find less invasive. And is reversible much quicker must less hassel so that that’s I felt slightly unethical because again it was money – but it was a wider system so money cost effectiveness so that ...“*

.....

In the first interview the GP got more into the theme of how the politics are creeping into both the recommendations and the targets – and stressed the lobbying activity:

GP:” *Well there are all sorts of lobbies I mean I used the mental health lobby which I am quite sympathetic to, but it’s it’s ahh a misinterpretation of what was originally intended and potentially difficult and dangerous one and that’s where I said you got to watch out for the politics. “*

I:” Yes exactly so this could really be a good system for getting politics into the medicine?”

GP: *“Yes it`s inevitably and there is a public health agenda which again has a conflict with the public health, the health of the population will conflict with the individuals needs in the consultation.”*

.....

In the second interview we came into discussing advises about the further development of QOF and politics where clearly lifted as an issue:

A: *“... separate from politics, because I a big anxiety that I have come across is that the targets are less driven by evidence and more driven by a particular party or government looking good or saying they got a good deal.”*

5. Discussion

5.1 The main findings

Positive effects:

Systematic approach and quality improvement

In both interviews the GPs expressed that QOF had lifted the quality in the GPs collegium and brought everyone to the same bar. They also appraised QOF for securing important measurements in a clinical stressed everyday practise.

In the literature I have found that measurements show improvement on several condition (hypertension, diabetic and asthma) – it`s uncertain whether this has a lasting effect, but the improvement for these conditions seems clear. The causality is though unclear. Is it due to QOF or is it due to a large-scale investment in quality programs from 1998 in UK (1)?

Quality

QOF seems to have lifted the quality especially for the doctors who did not perform well. But quality must be recognized as a broader concept and a broader aim than what`s covered in QOF. This is to be found both in the literature and in the interviews. The questions to what level quality is lifted and how lasting the effects are have to be asked.

Negative effects:

Patient centred care

Both in the literature and in the interviews it was problematized that the centre or the focus of the health care, is at risk to change or are changing. From meeting the patient open for his or her agenda and needs, QOF targets must be brought into the consultation and can push the patient`s agenda or need aside. The shift or displacement in primary care can be described as a shift from a patient centred care to a health service centred care.

Overtreatment

This issue is sparingly mentioned in the literature, but underlined in both interviews. The targets in QOF does not fit all and will be unreachable for many patients or may even be a threat for the patient and something the patient don`t want to reach for. The doctors feel obligated both to the system, to the patients and to their colleges. They can even feel forced

into chasing targets and that can sometimes be at the expense of the patients. It can give the patients reduced life quality (hypotension with dizziness, symptoms from hypoglycemia and so on) or even more dramatic consequences.

Political tool

The whole frame of QOF and the history in UK, tells us about the political pressure behind introduction of QOF. The QOF can also be, or is, a potent political tool. In both interviews this was addressed. The QOF system can be a political tool, by politicize the medical targets or guidelines.

Ethical issues

The authorities, society, patients and the doctors have to give the system reflections on how QOF influence the core values in family medicine and core value of health care in general. This is stated both in the literature and in the interviews.

In this report I have used both literature search and interviews as methods to get information about the P4P/QOF system.

5.2 Methodological considerations

The literature searches were not systematically. Searches were done with help from a librarian at the library of Faculty of medicine University in Oslo, by myself and I have also been tipped about some articles from colleges. Searches are described in Chapter 2, Methods and measurements.

By choosing a non-systematic search in literature, I did not achieve a systematic overview of or complete knowledge about the number of articles on P4P. I did not aim for a systematic review, and I consider the approach I used as suitable. As this thesis is based on qualitative methodology, I have all the time had some research question in mind when searching for articles, cf. chapter 2. I also considered; “what is the price for quality challenges resolved through a financial incentive system like P4P or QOF ?” , “what results have QOF given?”, “What are the effects on medical parameters?”, “What are the impact on the patient- doctor relationship?”, ”What are the effects on core values in the science of family medicine?” And “What ethical reflections or issues do literature and GPs have on P4P/QOF ?”. These questions among others, have been guiding the collection process, both when searching for literature and when talking to the interview participants and other colleges or stakeholders

here in Norway. I have done several searches and worked for sampling articles studying different consequences of QOF or P4P. Most of the articles are from UK, but also from other countries.

My supervisor and I decided it would be of more value for this thesis to supply and triangulate the information from literature with interviews, than - due to time – giving structural search the time disposable. I wanted to increase insight into P4P/QOF by also hearing GPs experience and thoughts about the incentive system.

The two interviews were planned for as two focus group interviews. The difficulties in recruit to the interviews turned the interviews into one depth interview and one mini-focus group, more about the challenge getting participants is found in chapter 2, Materials and Methods.

I had given different Qualitative sampling methods for interviews and focus groups, much thoughts before I started to recruit cf. chapter 2, Materials and Methods.

My method for recruitment for the two interviews was snowball sampling cf. chapter 2. Materials and method. Through this “snowball method” of recruiting I had no control over whether the participants would share views, disagree with each other, whether their opinion about QOF were typical or whether they had views different from most other GPs (deviant cases). Neither did I have any possibility to influence differences in age, sex, experience and area of London they worked. The limitation in my lack of flexibility to gather GPs from all over UK does sure also affect the selection of GPs.

One may criticize the recruitment method, but I think it was suitable. Different views are articulated in the material - both positive and negative aspects - and my informants was willing to share their views and experiences.

5.2.1 The two interviews

The participant in the depth interview was a male GP who had been working for about three decades as a GP. He had worked both before and after the introduction of QOF.

Many of his opinions on QOF were confirmed by the participants in interview number two.

In the second interview, the mini - focus group, there were four GPs with pretty similar characteristic; they were young, three qualified in 2008 as GP, the last qualified in August 2012. Three were female, one male. The four participants were salaried or freelance locums.

There were some different opinions and reflections on some subjects, but their main opinion about QOF I experienced were pretty similar and unison.

As the participants in total only count five GPs the sample is small and their experience not possible to generalise. At the same time their experiences are personal and reflect their view on the QOF and their experiences from the system they are working within.

Is the similarities in experience and reflections due to a very common opinion in the GP collegium or is this due to a biased recruitment to this study? Is the participants more active and more aware and reflected about the system than their college? And is this what make them to participate in this survey? Are they more positive to the QOF than their colleges or are they more critical?

I will not be able to answer these questions – but the use and interpretation of the information must be held against these perspectives.

5.2.2 Reflexivity

As I have said something about the reason for my curiosity about the P4P was that it provoked me, I was aware of my starting point, cf. chapter 1, Background. During the autumn 2011 I was quite sceptical to the quality and outcome framework. But reading literature and working with the analysis of the two interviews, I have earned more knowledge, and along with that a more balanced perception of the framework. Still sceptic but, more aware of many of the effects described in the articles about P4P and therefore moderating a bit my sceptic and got even more interested in understanding consequences of the arrangement of QOF.

I was aware of my attitude, before I went to London to do the interviews. I had thought through how to handle my prejudiced opinion in the interview situation. Some of the questions may have biased value, but as the interviews went through I learned by the way the GPs described both deficiencies in the system of QOF, side effects and even perverse effects, that everyone described after all a positive attitude. They “acclaimed” the framework for lifting everyone up to an acceptable level of quality delivery.

As said I have been aware of my thoughts about QOF all the time, I have discussed it with my fellow students and with my supervisor. And as I have learned a lot more about QOF trough my study, I have had to reconsider my thoughts and find me now, as said earlier, thinking more balanced about the incentive system, all though far from (only) positive.

Again and at the same time with consciousness about input from the interviews are five personal stories.

5.3 Thematic discussion:

5.3.1 A more systematic approach?

One of the main findings in the interviews was that QOF had reduced the variation in delivery in the GPs population. QOF had given the GP a systematic way of practise many important measurable issues and medical topics or problems. All the participants agreed on that QOF helps to bring everyone to the same bar. At the same time they problematized that after reaching “the level” it was a short way to get interrupted and disturbed by QOF (cf. 5.3.2). They said that QOF is not about a random screening, but it puts everyone to be checked on important measures like blood pressure, blood glucose and so on. This was what the participants in the interviews really put as the positive and awaited help and support QOF had given them in their everyday practice.

The GPs felt that QOF had helped them to structure and to remember many control parameters that could otherwise slip. In that way QOF helped them to become better doctors for their patients.

But is this due to QOF alone?

Financial incentives for clinicians are an intuitively reasonable solution to the well documented gaps between evidence based best practise and routine care (68). The need of good system and check lists can be one way to reach the goal of improved quality. It though implies that we believe that there is a causational connection between improved care and results for the patients due to change in the delivery of schematic measurements medical issues (68).

Many studies conclude that there is no reliable correlation between QOF and the more systematic approach in family medicine in UK today. There were already a huge emphasis on raising quality level in the GP surgeries by giving guidelines and decision support for the clinical practise before the introduction of QOF in 2004. May be this could have given the same result of structuration as QOF now is given the credit for?

Studies also show that the impact from QOF indicators and scores lose their effects after some time and flattens out (38). Can other systems for quality improvement be more robust over a longer time period?

Performance based pay may increase output for straightforward manual tasks. However, a growing body of evidence from behavioral economics and social psychology indicates that rewards can undermine motivation and worsen performance on complex cognitive tasks, especially when motivation is high to begin with (69).

The literature is based on lots and lots of data. The IT systems`collection of data is huge and give an imaginary base for performing studies and research. There has also been a lot of research on this data. These studies show that some medical conditions seem to be well suited for arrangements like QOF. Blood pressure/hypertension, asthma and diabetic are good examples. It seems like these conditions have been better regulated, but again there is no possibility to establish direct causality between QOF and improved care as long as many other actions have been taken, cf. Background history England, chapter 1.

5.3.2. Patient centred health care

In the interviews all the participants address the conflict between patient centred care and the realities of the agenda that has been set for them by QOF. The QOF targets and boxes to tick interferes with the consultation and not rare with the patients agenda. To their experience it moves the focus away from the patient toward the QOF targets and the “doctor`s agenda” and computers demand for ticking boxses. They also problematize been pushed “to wear” an agenda by having to fulfil or meet the targets of QOF. In the literature this is very little problematized, and to my opinion that`s a really huge gap. This need to be addressed, this need to be discussed, and this need to be studied. Which focus do we want to choose for our health service? This has to and will be regulated by the arrangement and structure of health care systems. QOF is an important arrangement in the way it intrude with the contact between patient and the doctor. This is also further discussed under section 5.Etichs, in this chapter.

How can QOF measurements be oriented and arranged so that the interference of the patient – doctor relationship reduces? The description from the interview participants of ticking boxses or “shoe horning in” obligated QOF schemes is problematic. And it gets even more problematic the more that`s put into QOF. The more indicators and the more schemes or obligated guides put into QOF, the more the doctors have to secure that they deliver on demanded areas regardless of the patient`s agenda and needs.

I think the reason why the interview participants - and the literature- not are giving thumbs down for the QOF in total, is because it brings every GP to the same bar of quality and, after all, make only 20% of the income. It is possible to ignore the QOF demands, when they

disturbs too much or are evaluated unethical and the systematic effects justify many of the side effects or even perverse effects.

5.3.3 Overtreatment

Another important issue addressed in the interviews, but little in the literature, is overtreatment of the patients due to reaching the best score of QOF.

“one size doesn’t fit all” This statement holds much and it’s not only an important statement for the medical optimal targets, but it is also an important statement for the essence of the medical science – to treat every person individual as the unique medical human being it is.

All interview participants addressed that QOF can contribute to reduce the patient’s life quality, because they can be treated “too good” or “too hard” to reach the set target in QOF. They addressed the conflict as the dilemma when the holistic medical examination or assessment done, gave them reason to recommend the patient to follow other optimal personalized targets than them set by QOF. They – and especially the younger doctors felt pressured to meet QOF targets. In both interviews the statement “it might also kill them” was given. The latter I have not seen described in the literature, so I will focus on the statement of reduced life quality. The QOF targets are made to fit all – every doctor know that many patients will fall outside the optimal targets. The variation among individuals are huge, both in personalities, in history, in life experiences, in capacity physically and mentally and so on. This variation can also be found in physical parameters due to genetics, sex, age, weight and ethnicity and many other things. How can we reduce the side effects when we want to make one quality systems fit everyone? I don’t think there is any other answer than making the system as flexible as possible.

May be the further development of QOF have to make the framework more flexible for individual adjustments, but is that possible without losing the systematic?

5.3.4 Political tool

The change of incentive system in UK in 2004 was given two reasons: international studies telling about too low quality in UK health care and the political statement saying England did spend less money in health care compared to other comparable countries (3). The first was already addressed through NHS widespread program of quality improvement under the general heading of “clinical governance” (1). There were many initiatives under this program and the quality had increased in health delivery. So there may be a good reason to ask if the main reason for implementing QOF and changing the incentive system, was to show political

decisiveness? At least, as the quality already was about to rise, the argument about UK spending less money in health care compared to comparable countries may be the essential one? I think that it allows to say that the introduction of the incentive system, QOF was politically motivated.

But it's not only history that puts QOF in a political frame, more important it is described as a potent political tool; "a way of the authorities to intrude into every consultation."

This is found little about in the literature, may be not surprisingly delicate as it can be. But it was addressed and stressed in both interviews. They addressed lobbies and that by lobbying both important medical issues can be included into QOF, but also problematized when used and misused by medical industries or political/financial motives alone. QOF is described as a suitable system to get political targets into medicine. The worries of getting the public health agenda's conflicting with the public health (health of the whole population) and one or both of them set up against the individual needs in the consultation.

In both interviews they told me about the third month injection for prevent unwanted pregnancies - and they all felt that this guideline was very political and not founded on medical knowledge.

Is this only negative, or does it give the politicians – and the doctors a helping hand in the difficult and challenging matter of prioritization? Can it be defended to incorporate directions/instructions about preferred treatment in guidelines, QOF targets and medical practice due to political and economic arguments?

I think yes – and I see that there is a real opportunity to get into the core of prioritization with an incentive system infiltrating the clinical delivery.

But and there is a big **but**; because this can only be on one condition – an absolutely crucial condition:

The politicians have to put their personal political agendas and personal political interests aside and base the system on - as far as possible objective criteria of severity, effect and cost effectiveness. A wide understanding⁹ of EBM must be the basis, and the lack of EBM in primary care have to be taken seriously and given a considerable investment to get enough research on the unselected patient population, cf. 5.3.5.

⁹ With wide understanding we have to include experience-based knowledge.

And this condition may be not achievable before the politicians and humans become angels...

To my opinion the real danger having a framework and incentive system that can be such a potent political tool, is that the medical science is politicised and may, in extreme situations, be put at risk to no longer serve the patient, but then serves the politicians, or even capitalists (as they are or are trying to influence the politicians through lobbying). The risk is huge that if/when/as politics are creeping into both the recommendations and the targets, the medical science and evidence based medicine are put aside.

5.3.5 Quality

The definitions of quality are many and the literature emphasise different dimensions. Some of the definitions are presented in chapter 1. Background. IOM's definition is widely used: Quality is the "extent to which health services provided to individuals and patient populations improve desired health outcomes." (13)

Quality of care must be seen in the context of the health care system. This is important in the evaluation of P4P and other financial system – like Fee For Service (FFS), capitation a.o. Communication is the most important aspect of context, and evidence shows that more important than the length of the consultations is the quality of the doctor- patient relationship. This relationship is also important in aspect to adherence to therapy and even outcome (14). The policy evidence takes in efficiency and equity. Due to scarce health resources the importance of economic analysis are increasing and are therefore taken into the planning for health care intervention and innovations.

The increasing availability of medical evidence in clinical practice was expected to improve the quality of care. It has been heavily invested in EBM and the body of evidence is increasing rapidly. An important restriction of the practical usefulness of medical evidence is that most research is not driven by clinical relevance, but rather by commercial interest (14) As an example there are many studies and vast number of publications on pharmacological treatment, but only a few studies are published on the effect of exercise and lifestyle changes (14). In my opinion this also give reason for concern and question must be asked to what extend the research dimension of EBMs is biased? And there is also a danger that EBM pursues what is possible, rather than what is relevant (14).

Another challenge is that the evidence of medicine is increasing, but as said in chapter 1, we have to take action on the major flaws and gaps hampering its applicability in primary care. One of the reasons for that, as in Norway for example, are the huge challenge due to almost every penny for research goes to hospitals or secondary care, while primary care, and for sure family medicine, get very little attention and very small amount of resources. The patient population in primary care is not collated and therefore quite different from the patient population in secondary care. This has to be addressed if we are to get quality improvement to be based on EBM in primary care, as also addressed in 5.3.4, Political tool.

Just as a clear and uncontested definition of quality in primary care are unrealistic, there are some specific component that are widely agreed on as central to the idea of quality (17).

Barbra Starfield has described these central ideas as four unique features of a primary care service; access, person focused care over time, comprehensiveness and coordination. (17).

“There is lots of evidence that a good relationship with a freely chosen primary-care doctor, preferably over several years, is associated with better care, more appropriate care, better health, and much lower health costs. In contrast, little is known about most of the benefits of specialty care, although we do know that the greater the supply of specialists, the greater the rates of visits to specialists. We also know that when specialists care for problems outside their main area of expertise, the results are not as good as with primary care. Since most people with health problems have more than one ailment, it makes sense to have a primary-care practitioner who can help decide when specialist care is appropriate.” Barbra Starfield

Barbra Starfield was an accomplished researcher and champion of the need for a strong primary care system in the U.S. and worldwide. Her work led to the development of important methodological tools for assessing diagnosed morbidity burden, including the Adjusted Clinical Groups (ACG) System, the Primary Care Assessment Tools and the CHIP tools for assessing adolescent and child health status. She has been described as” a giant in the field of primary care and health policy...” by Michael J. Klag, MD, MPH, dean of the Johns Hopkins Bloomberg School of Public Health¹⁰.

¹⁰ <http://www.jhsph.edu/news/news-releases/2011/starfield.html>

With this context or frame for the family care's importance as a corner stone or girder of the health care system it should be naturally paid much attention to what will strengthen the family medicine and what will not.

Is P4P or QOF a way of strengthening primary care? Does it improve access to the GPs, strengthen the person focused care over time, comprehensiveness and coordination? To my opinion there are no yes or no answers to these questions, but as the discussion and reflections show in this chapter and there are some positive aspects and there are some negative. I think a major deficiency is the lack of consequence analysis and piloting performed before UK introduced a nationally framework of such an extent. Without knowledge about the impact and results of QOF on family medicine it really was a huge national pilot and it has proved both side effects, perverse effects although also positive effects.

When it comes to the question of whether QOF has raised the level of quality in primary care - the answer will depend on the definition of the term and which dimensions of quality address. I will question to what extent it has strenghten and facilitated the ethical dimensions in primary care – the patient- doctor relationship and the respect for the individual human in need for or seeking help with the doctor, cf. 5.3.6. It seems to have given more systematic and raised the quality of some diagnoses. The next question is – but to what level? Only to the level of justifiability?

In the second interview the participants gave me the expression that QOF only secured quality to a certain level and when reached that level, the QOF targets and registrations start to interfere with your practise.

This is also supported in the literature. The positive effects were higher for initially low performers compared to already high performers (22).

Although this worrying about the interference from the QOF system, both participants in the first and second interview, addressed the boxses popping up in their IT system(Electronic Medical Journal/records, EMJ/EMR) as a good help. It helps them not forget important medical measures.

In the literature the question about whether the quality already was improving due to NHS widespread program and quality improvement in 1998 under the heading of “clinical governance” or in fact the quality improvement was due to introduction of QOF is discussed. The answer is not clear. Some diseases or conditions targeted in QOF seems to have got

improved, like asthma and diabetes (36), but the improvement seems not lasting (36). Other studies conclude that making information on performance available for a wider range of conditions is an essential component of quality improvement (65) because it will give a general positive effect to quality improvement.

In the first interview the experience GP says that quality is improved in a clinical view in relation to medical conditions with hard end points, where the targets in QOF are thoroughly described with evidence.

Do they have access to patient data through out some years to compare the development? And if they do, do they have the capacity to check and evaluate whether the measurements and the quality is lowering after two or three years?

Or is the literature deficient in catching or display that the patient may be even better taken care of in a holistic view, when some QOF targets go down, or flattens out, as the total care may improving?

There are lots and lots of data – because all the registrations done by the GPs by ticking boxses are collected and can give the authorities, patients and the doctors information. But there are still many questions about causation? What are the effects on quality due to targets in QOF and due to the system of QOF? What will be the best way of developing the system to constant improve quality? Is it to include more targets and more diseases into QOF? Or will that undermine the quality improvement?

QOF can be a contribution to quality improvement or may be better; can be a contribution to quality securement, but as pointed at earlier, it will never cover all dimensions of quality and it's important to never use it as the only tool.

I will get back to the definition of quality. I think it is very important to not get lost or blinded by the available policy instruments precented by QOF. Quality definition can't be covered by QOF indicators or guidelines alone. Quality is a broader concept and must be described by several aims as in the IOM and Norwegian definition. Aims that all will contribute to development of quality in practise and the clinical every day.

Safe - Avoiding preventable injuries, reducing medical errors,

Effective - Providing services based on scientific knowledge (clinical guidelines),

Patient centered- Care that is respectful and responsive to individuals

Efficient - Avoiding wasting time and other resources

Timely- Reducing wait times, improving the practice flow and

Equitable- Consistent care regardless of patient characteristics and demographics¹¹

5.3.6 Ethics

Medicine is a value based practise. Every consultation, every patient - doctor meetings are connected to value choices, and therefore medicine gives the doctor a responsibility to manage this wise and with respect and always in the best interest of the patient. At the same time the GP have a corporate social responsibility. The GPs are put to administrate huge resources on behalf of the society. How do these two missions go along? The patient has the right to get access to an open, predictable and unprejudiced health care and the doctors agenda, personal believes or personal tolerance for differences are not supposed to interfere with the consultation or the doctors medical judgment of patient`s needs.

How do a system with financial incentives fit into this?

How should the doctor meet demands and instructions from the authorities? Are they obligated through their contract as GP to do what they are told from the employer (authorities) or are the doctors obligated to be faithful to their medical ethical principles (autonomy, utility and justice), recognising that individual patients' views on treatment and autonomy may be in conflict with quality improvement (QI) schemes(60).

In one article (62) it concludes that current arrangements (P4P) are based on fundamentally acceptable ethical principles, but are guided by an incomplete understanding of health-care quality. Furthermore, their implementation without evidence of safety and efficacy, is ethically precarious because of potential risks to stakeholders, especially vulnerable patients.

Buetow and Entwistle (61) conclude in their reflections about pay-for-virtue as an option to improve pay for performance, that concerns about current P4P schemes eroding professional values and undermine moral motivation might be ameliorated to some extent by 1) investment in pre-qualification moral training and post-qualification support for ongoing moral reflection

¹¹ http://www.peerpt.com/website/index.php?option=com_content&view=article&id=10&Itemid=10

and development.2) Modification to current pay-for-performance schemes to improve the assessment of individual patient care and/or 3) the addition of indicators of virtuous practice.

Snyder et al (63) says in 2007 in their ethics manifesto that by framing the discussion of P4P in terms of ethics and professionalism both in the context of the individual patient and for society, they hope to move the debate forward with a patient centered focus - one that puts the needs and interests of the patient first—as P4P programs evolve (63). They continues that P4P programs may have the potential to increase overall quality of care when aligned with the ethical obligations of the physician to deliver the best quality care to her or his patient, but the risks are that incentives could result in de-selection of patients, “playing to the measures” rather than focusing on the patient as a whole, misalignment of perceptions between physicians and patients, and increased unnecessary care and costs, have the potential to harm access to care, continuity of care, patient-physician relationships, and care for those patients with complex chronic conditions (63).

In the interviews the participants addressed ethical dilemmas in different ways. In the first interview the ethical dilemma was addressed due to in what way the QOF intrude into the patient`s life, putting or pushing the patients need away and forcing the doctor to give advises or give treatment due to the measurements and demands in the QOF rather than a holistic medical approach based on the patient need. The younger doctors addresses the ethical dilemmas also due to driving the focus away from patient centred care, and they went more into details of the system and the way they manage to handle the different demands – whether it was ticking the boxes or filling out schemes into a consultation on a false or honest base. In both interview they addressed and expressed concerns about reduced empathy with the patient due to the pressure of fulfilling the QOF targets.

In the second interview they also, interesting, addressed the link between QOF score/ticking boxses and the money as an ethical issue. And they raise the question whether it would be easier to defend the QOF demands if they were not connected with money? All the GPs in the second interview were salaried or freelance doctors. May be they are more free to reflect on this issue than the partner GPs are? The partner GPs could may be so “infiltrated” with the economy that there are no longer space or possibility to do the reflection without going out of the partnership? This was hinted in the second interview, but as this was the second and last interview I could not get this confirmed or denied by the first GP directly. The first GP was a partner. The participant in the first interview answered the question whether he encountered

any ethical dilemmas due to QOF with several views, and he also addressed the economy by saying that when he found himself thinking about his money rather than what the patient wants or even what is best for the patient, that sure is an ethical dilemma.

The participants in the second interview also find the QOF as an ethical support to their work, giving them a security to give a good medical standard on their service.

Using clinical audit for financial reward and punishment, rather than in a collegial and reflective effort to improve care, amplifies the challenges of performance measurement. Incentives may mutate honesty into legal trickery; gaming can so thoroughly distort reality that rewards become uncoupled from performance.(69)

I'm left with the important question;

Is there a reason for us to be worried that pay for performance may not work simply because it changes the mindset needed for good doctoring(69).

6. Conclusion

Our future health care system must accommodate new requirements concerning quality and patient safety, and should be based on the ethical mindset and platform of good doctoring and good nursing. This frame must be able to cover the complexity and even increasing complexity found in the future health care. With the development of medical knowledge and the medical technology, we will face more complexity into the individual level, microsystem level (patient-doctor relation), at the system- and the society level. We will face more multi-morbidity with the demographic changes where people live longer with more diseases. The heterogeneity of the patient population will grow and the social inequalities in health, are in danger of being even bigger. We will be faced with the informed and resourceful patient who follows up his or her own treatment, and we will have the poor or less resourceful patient population not able to take that responsibility or demanding the system for help.

Financial incentives for clinicians are an intuitively reasonable solution to the well documented gaps between evidence based best practise and routine care (68). The need of good system and check lists can be one way to reach the goal of improved quality. It though implies that we believe there is a causational connection between improved care and results for the patients due to change in the delivery of schematic measurements medical issues. From the literature the conclusion is clear; although huge amount of data, the evidence on the efficiency of P4P is scarce and inconclusive. The side effects of implementing P4P are many and have to be more thoroughly studied.

In order to improve quality we have to include all the dimensions of the term; **Safety, Effective, Patient centered, Efficient, Timely, Equitable**¹².

Let me give a description;

In primary Care the quality can't be measured in one diagnose or condition alone. We live our life in relations and complex contexts. As a doctor we have to meet the whole person, not treat a diagnosis or blood pressure alone. There are yet no tools capable of assessing the quality of primary care delivered, to those who have multiple and compounding conditions. Such tools will be dependent on a much richer understanding of the complex interactions between

¹² http://www.peerpt.com/website/index.php?option=com_content&view=article&id=10&Itemid=10

different conditions, diagnoses, and contexts and of the challenges of prioritizing, integrating, and personalizing care for a succession of different individuals, families, and communities (70). The system of P4P seems to have eliminated the worst practices, but at the same time it leaves little room and no incentives for the exploration of excellence. Attention confined to the parts may damage the whole. For example, focusing on the treatment of blood pressure may improve this indicator, but does not necessarily reduce patients' global cardiovascular risk. The future must lie in getting balanced and well tested combination of different quality improvement systems and incentive systems.

Future attempts to measure quality in primary care should take the unique complexity of primary care into account and move away from a linear approach to engage multiple perspectives and multiple levels (Substructure literature F).

Incentive systems have no value in themselves; they are just one among many tools to influence what happens in clinical practice. On the other hand the P4P seems to have brought every GP to the same bar, and by that may have contributed to standardisation and more equitable service in primary care. To find the optimal incentive system, I think we have to combine tools and approaches. Some clear medical condition can be regulated and get improved focus by give special incentives to get the systematic into the treatment. If we manage to combine P4P with per capita(capitation) and FFS and may be also PP4P(71) we will be closer to an optimal economic regulation and support for the health care. The question we are left with are if it will be cost effective, because this will always be just one tool. Along with this or even more important and potent, to my opinion, we have to develop systems for reflection. Systems where doctors and surgeries can come together and find ways to improve their quality, define their challenges and bench mark their improvement. This will ease doctors and surgeries systematic work with quality development and patient safety locally. We have to have systems that facilitate a culture of quality improvement.

With a combination of systems for reflection that supports a systematic quality approach to the health service and patient safety, along with a mix of incentive system, I think we will be closer to develop in a good and safer way the future medical mindset to conserve the for ever core of health care;

To heal sometimes, to alleviate often and always comfort.

7. Reference list:

1. Roland M, Gunthrie B, Thome DC. *Primary medical care in the United Kingdom. J Am Board fam Med* 2012 Mar;25 Suppl 1:S6-11
2. Roland M. *Linking physician pay to quality of care: a major experiment in the UK. N Engl J Med* 2004;351:1448-54.
3. Iversen T. Prestasjonsavhengig betaling til legene - noe for Norge? RØST: Helse 2008, nr. 1. www.okonominettverket.no/Artikkel/804.html (26.6.2012.)
4. Steel N, Maisey S, Clark A et al. *Quality of clinical primary care and targeted incentive payments: an observatorial study ,BMJ* 2007;449-454
5. Whalley D, Gravelle H, Sibbald B. *Effects of the new contract on GPs` working lives and perceptions of quality of care: a longitudinal survey. Br J Gen Pract.* 2008 Jan;58(546):8-14.
6. Doran T, Fullwood C, Gravelle H, Reeves D, Kontopantelis E, Hieroech U and Roland M. *Pay-for-Performance in Family Practises in the United Kingdom. N Engl J Med* 2006;355:375-384 July 27, 2006
7. Munkerud S. *Økonomiske incentiver og behandlingskvalitet. Tidsskriftet* 2010;130:2030-2
8. Melberg HO, Kjekshus LE. *Fremtidens Helse-Norge s. 187-202. Fagboklaget* 2012.
9. Doran T, Kontopantelis E, Fullwood C et al. *Exempting dissenting patients from pay for performance schemes: retrospective analyses of exeption reporting in the UK Qulaity and Outcomes Framework. BMJ* 2012;344:e2405 doi:10.1136/bmj2405
10. Doran T, Fullwood C, Reeves D et al. *Exclusion of patient from pay-for-performance targets by English physicians. N Engl J Med* 2008 Jul 17;359(3):274-84
11. National Primary care, Research and Development Centre. *Using The GP Patient Survey To Improve Patient Care: a guide for general practices.*
12. Roland M, Elliott M, Lyratzopoulos G, et al. *Reliability of patient responses in pay for performance schemes: analysis of national General Practitioner Patient Survey data in England. BMJ* 2009;339:b3851.
13. http://www.peerpt.com/website/index.php?option=com_content&view=article&id=10&Itemid=10
14. Van Driel ML, De Sutter AI et al. *Quality of care: the need for medical, contextual and policy evidence in primary care. Journal of Evaluation in Clinical Practise,* 11,5,417-429.
15. Donabedian A, *The quality of care , how can it be assessed? JAMA, SEP 23/30, 1988*
16. Mold J, Blake G and Beeker L. *Goal-oriented medcal care. Family Medicine* 1991 23,46-51
17. Peckham S, Wallace BA. *Pay for performance scheme in primary care: what have we learnt ? Quality in Primary care* 2010;18:111-16.
18. Ayelet Kuper, 1 Lorelei Lingard, 2 Wendy Levinson 3. *QUALITATIVE RESEARCH Critically appraising qualitative research. 688 BMJ | 20 SEPTEMBER 2008*
19. Tjora A. *Kvalitative forskningsmetoder i praksis. Gyldendal 1. utgave, 1. opplag* 2010
20. Witter s, Fretheim a, Kessey FL, Lindahl AK. *Paying for performance to improve the delivery of health interventions in low- and middle –income countries. Cochrane Database Syst. Rev.* 2012 Feb 15;2:CD007899
21. Emmert M, Eijkenaar F, Kemter H et al. *Economic evaluation of pay-for-performance in health care: a systematic review.*
22. Van Herck P, De Smedt D, Annemans L et al. *Systematic review : effects, design choises, and context of pay- for- performance in health care. BMC Health Servies Reseach* 2010;10:247
23. McDonald R, Roland M. *Pay for performance in primary care in England and california: comparison of unintended consequences. Ann Fam Med.* 2009 Mar-Apr;7(2):121-7.
24. Bell CM, Levinson W. *Pay for performance: learning about quality. CMAJ* 2007 June 5;176(12):1717-1719.

25. Chauvel N, Le Valliant M, Pelletier-Fleury N. *Variation in HbA1c prescription for patient with diabetes in French general practice: an observational study prior to the implementation of a P4P programme.* *Eur J Public Health.* 2012 Jan 4.
26. Crosson J, Ohman-Strickland PA, Campbell S, Phillips RL et al. *A comparison of chronic illness care in US and UK family medicine practices prior to pay-for-performance initiatives.* *Family Practise* 2009; 26:510-516.
27. Simpson CR, Hannaford PC, Ritchie LD, Sheikh A, Williams D. *Impact of the pay-for-performance contract and the management of hypertension in Scottish primary care: a 6-years population-based repeated cross-sectional study.* *BR J Gen Pract.* 2011 Jul;61(588):e443-51.
28. Kirschner K, Braspenning J, Akkermans RP et al. *Assessment of a pay-for-performance program in primary care designed by target users.* *Family practise*(2012)doi:10.1093/fampra/cms055
29. Casalino LP, Alexander GC, Jin L, Konetzka RT. *General internist view on pay-for-performance and public reporting of quality scores: A national survey.* *Health Affairs*, 26 no.2(2007):492-499.
30. Millett C, Netuveli G, Saxena S, Majeed A. *Impact of pay for performance on ethnic disparities in intermediate outcomes for diabetes: a longitudinal study.* *Diabetes care* 2009 Mar;32(3):404-9. Epub 2008 Dec 10.
31. Millett C, Gray J, Saxena S, Netuveli G, Khunti K, Majeed A. *Ethnic disparities in diabetes management and pay-for-performance in the UK: the Wandsworth Prospective Diabetes Study.* *PLoS Med.* 2007 Jun;4(6):e191
32. Vaghela P, Ashworth M, Schofield P, Gulliford M. *Population Intermediate Outcomes of Diabetes under Pay-for-Performance Incentives in England From 2004 to 2008.* *Diabetes Care.* 2009 March;32(3):427-429.
33. Millett C, Gray J, Saxena S, Netuveli G, Majeed A. *Impact of a pay-for-performance incentive on support for smoking cessation and on smoking prevalence among people with diabetes.* *CMAJ.* 2007 Jun 5;176(12):1705-10.
34. Campbell S, Reeves D, Kontopantelis E et al. *Quality of Primary Care in England with the Introduction of Pay for Performance* *N Engl J Med* 2007; 357:181-190 July 12, 2007
35. Ashworth M, Medina J, Morgan M. *Effect of social deprivation on blood pressure monitoring and control in England: a survey of data from the quality and outcome framework.* *BMJ* 2008; 337doi:10.1136/bmj.a2030 (published October 2008)
36. Serumaga B, Ross-Degnan D, Avery AJ, Elliot RA et al. *Effect of pay for performance on the management and outcomes of hypertension in the United Kingdom: interrupted time series study.* *BMJ* 2011; 342doi: 10.1136/bmj.d108 (Published 25 January 2011) Cite this as: *BMJ* 2011; 342: d108
37. Lee JT, Netuveli G, Majeed A, Millett C. *The effects of pay for performance on disparities in stroke, hypertension, and coronary heart disease management: interrupted time series study.* *PLoS One.* 2011;6(12):e27236. Epub 2011 Dec 15.
38. Stephen M. Campbell, Ph.D., David Reeves, Ph.D., Evangelos Kontopantelis, Ph.D., Bonnie Sibbald, Ph.D., and Martin Roland, D.M. *Effects of Pay for Performance on the Quality of Primary Care in England.* *N Engl J Med* 2009; 361:368-378 July 23, 2009.
39. Chien AT, Wroblewski K, Damberg C, Williams TR et al. *Do physician organizations located in lower socioeconomic status areas score lower on pay-for-performance measures?* *J Gen Intern med.* 2012 May;27(5):548-54. Epub 2011 Dec 13.
40. Doran T, Fullwood C, Kontopantelis E, Reeves D. *Effects of financial incentives on inequalities in the delivery of primary clinical care in England: analysis of clinical activity indicators for the quality and outcomes framework.* *Lancet* 2008; 372: 728-36. Published online August 12, 2008.
41. Fleetcroft R, Parekh-Bhurke S, Howe A, Cookson R, Swift L, Steel N. *The UK pay-for-performance programme in primary care: estimation of population mortality reduction.* *Br J Gen Pract.* 2010 Sep;60(578):e345-52.

42. Campbell SM, Kontopantelis E, Reeves D, Valderas JM, Gaehtl E, et al. *Changes in Patient Experiences of Primary Care During Health Service Reforms in England Between 2003 and 2007. The Annals of Family Medicine* November 1, 2010 vol.8 no. 6 499-506.
43. Morgan CL, Beerstecher HJ. *Satisfaction, demand, and opening hours in primary care: an observational study. Br J Gen Pract.* 2011 Aug;61(589):e498-507.
44. Addink RW, Bankart MJ, Murtagh GM and Baker R. *Limited impact on patient experience of access of a pay for performance scheme in England in the first year. Eur J Gen Pract.* 2011 Jun;17(2):81-6. Epub 2011 Feb 8.
45. Boeckxstaens P, Smedt DD, Maeseneer JD, Annemans L, Willems S. *The equity dimension in evaluations of the quality and outcomes framework: a systematic review. BMC Health Serv Res.* 2011 Aug 31;11:209.
46. Vamos EP, Pape UJ, Bottle A, Hamilton FL, Curcin V et al. *Association of practice size and pay-for-performance incentives with the quality of diabetes management in primary care. CMAJ.* 2011 Sep 6;183(12):E809-16. Epub 2011 aug 2.
47. Doran T, Campbell S, Fullwood C, Kontopantelis E, Roland M. *Performance of small general practices under the UK's Quality and Outcomes Framework. BR J Gen Pract.* 2010 Sep;60(578):e335-44.
48. Checkland K, Harrison S. *The impact of the Quality and Outcomes framework on practice organisation and service delivery: summary of evidence from two qualitative studies. Qual Prim Care.* 2010;18(2):139-46.
49. Alyahya M. *Changing organizational structure and organizational memory in primary care practices: a qualitative interview study. Health Serv Manage Res.* 2012 Feb;25(1):35-40.
50. Maisey S, Steel N, Marsh R, Gillam S, Fleetcroft R, Howe A. *Effects of payment for performance in primary care: qualitative interview study. J Health Serv Res Policy.* 2008 Jul;13(3):133-9.
51. Campbell SM, McDonalds R, Lester H. *The experience of pay for performance in English family practice: a qualitative study. Ann Fam Med.* 2008 May-Jun;6(3):228-34.
52. McDonald R, Harrison S, Checkland K, Campbell SM, Roland M. *Impact of financial incentives on clinical autonomy and internal motivation in primary care: ethnographic study. BMJ* 2007 Jun 30;334(7608):1357. Epub 2007 Jun 19.
53. Checkland K, Harrison S, McDonald R, Grant S, Campbell S, Guthrie B. *Biomedicine, holism and general medical practice: responses to the 2004 general Practitioner contract. Sociol Health Illn.* 2008 Jul;30(5):788-803. Epub 2008 Apr 28.
54. Gemmel I, Campbell S, Hann M, Sibbald B. *Assessing workload in general practice in England before and after the introduction of the pay-for-performance contract. J ADV NURS.* 2009 Mar;65(3):509-15.
55. McDonald R, Harrison S, Checkland K. *Incentives and control in primary health care: findings from English pay-for-performance case studies. Journal of health organization and management, Vol.22Iss:1 pp.48-62.* 2008
56. Dalton AR, Alshamsan R, Majeed A, Millett C. *Exclusion of patients from quality measurement of diabetes care in the UK pay-for-performance programme. Diabet med.* 2011 May;28(5):525-31.
57. Campbell SM, Kontopantelis E, Hannon K, Bruke M, Barber A and Lester H. *Framework and indicator testing protocol for developing and piloting quality indicators for the UK quality and outcomes framework. BMC Family Practise* 2011;12:85.
58. Lester HE, Hannon KL, Campbell SM. *Identifying unintended consequences of quality indicators: a qualitative study. BMJ Qual Saf.* 2011 Dec;20(12):1057-61. Epub 2011 Jun 21.
59. Hannon KL, Lester HE, Campbell SM. *Recording patient preferences for end-of-life care as an incentivized quality indicator. What do general practice staff think?. Palliat Med.* 2012 Jun;26(4):336-41. Epub 2011 Jun 16.
60. Tapp L, Edwards A, Elwyn G, Holm S, Eriksson T. *Quality improvement in general practice: enabling general practitioners to judge ethical dilemmas. J Med Ethics* 2010;36:184-188 doi:10.1136/jme.2009.032078

61. Buetow S, Entwistle V. *Pay-for-virtue: an option to improve pay-for-performance?*. *Journal of Evaluation in Clinical Practice* Volume 17, Issue 5, pages 894–898, October 2011 Article first published online: 28 JUL 2011
62. Wharam JF., Paasche-Orlow MK, Farber NJ, Sinsky C, Rucker L, Rask KJ, et al. *High Quality Care and Ethical Pay-for-Performance: A Society of General Internal Medicine Policy Analysis*. *J Gen Intern Med*. 2009 July; 24(7): 854–859. Published online 2009 March 18. doi: 10.1007/s11606-009-0947-3 PMID: PMC2695523
63. Snyder L, Neubauer RL. *Pay-for-performance principles that promote patient-centered care: an ethics manifesto*. *Ann Intern Med*. 2007; 147792-4 PubMed.
64. Nelson A.R, *Pay-for-performance programs: Ethical questions and unintended consequences*. *The journal of family medicine*, Oct. 2007, Vol 56, No 10A.
65. Steel N, Bachmann M, Maisey S, Shekelle P et al. *Self reported receipt of care consistent with 32 quality indicators: national population survey of adults aged 50 or more in England* *BMJ*2008;337doi: 10.1136/bmj.a957(Published 14 August 2008)Cite this s:BMJ2008;337:a957
66. Sutton M, Ross E, Guthrie B and Watt Graham. *Record rewards: the effects of targeted quality incentives on the recording of risk factors by primary care providers*. *Health Economics* Volume 19, Issue 1, pages 1–13, January 2010
67. Walker S, Claxton K, Cookson R et al. *Value for money and the quality and Outcome framework in primary care i UK NHS*. *BR J Gen Pract* 2010 May 1;60(574):e213-e220.
68. Glasziou PP, Buchan H, Del Mar C, et al. *When financial incentives do more good than harm: a checklist*. *BMJ*2012;345doi:10.1136/bmj.e5047
69. Ariely D, Himmelstein D U. *Why pay for performance may be incompatible with quality improvement*. *BMJ*2012;345doi:10.1136/bmj.e5015
70. Fortin M, Soubhi H, Hudon C, Bayliss EA, van den Akker M. *Multimorbidity's many challenges*. *BMJ*2007;334:1016-7.
71. Wu J. *Rewarding healthy behaviors—pay patients for performance*. *Ann fam. Med*. 2012 May;10(3):261-3

Substructure literature

- A. Shi L, Macinko J, Starfield B, Wulu J, Regan J and Politzer R. *The relationship between primary care ,income inequality and mortality in US States. 1980-1995*. *Journal of the American Board of Family Practise* 2003 16, 412-422.
- B. Nederlandse Zorgautoriteit. *An optimal remuneration system for general practitioners 2007, research paper nr. 4*. www.nza.nl/publicaties/Researchpapers/ (28.06.2012).
- C. Suleman J, Anwaar MS, Weston C, Baker R. *Use of outcomes in monitoring healthcare- how many outcome measures are needed in monitoring diabetes in primary care?*. *J R Soc Med*. 2011 Oct;104(10):413-20.
- D. Starfield B, Mangin D. *An international persepective on the basis for payment for performance*. *Qual Prim Care*. 2010;18(6):339-404.
- E. Landon BE. *Use of quality indicators in patient care: a senior primary care physician trying to take good care of his patients*. *JAMA*. 2012 Mar 7;307(9):956-64. Epub 2012 jan 31.

F. Heath I, Rubinstein A, Stange K.C et al. Quality in primary health care: a multidimensional approach to complexity. *BMJ* 2009;338doi: <http://dx.doi.org/10.1136/bmj.b1242> (Published 2 April 2009)

G. National Primary care, *Research and development Centre*. *Using The GP Patient Survey To Improve, Patient care; a guide for general practise*.

H. Department of health. *Improving health and care: the role of the outcome frameworks*. 22 november 2012

I. Misselbrock D. *What is the right medicine? Practise, summer 2010*.

8. Appendix

Appendix 1

Problemstilling sent til bibliotekar før søk:

Masteroppgave: Pay for performance kvalitet til hvilken pris ?

Jeg ønsker gjerne å få hjelp til å søke på artikler som tar for seg

Hvordan har Pay for performance påvirket kvaliteten i allmennmedisin ?

Hvordan har Pay for performance påvirket fagutøvelsen i allmennmedisin ?

Jeg tenker særlig å se til England i min oppgave i det vi har samme

finansiering av våre helsesystemer, men er også interessert hvis det

finnes gode artikler fra andre land.

Nøkkelord på norsk		Allmenn-medisin	Kvalitet	Fagutøvelse	England		
Nøkkelord på engelsk	Pay for performance				United Kingdom? Great Britain		
Synonymer engelsk	P4p UK quality and outcomes framework?						
MeSH	Fee for service plans Reimbursement mechanichs Physician Incentive plans	General practice Family practice	Quality of health care Outcome and process assessment (Health care)	Professional practice?	England		
Emneord EMBASE	Outcome assessment Scoring system		Health care quality	Professional competence			

Databases:

- PubMed (MEDLINE)
- EMBASE
- Web of Science

Appendix 2

1. Before we start to hear about your experience with the QOF I would like us to take a very short presentation:

What's your age, how long you have been working as GP, are you specialist and where you are working. What's the size of your practice, are you a locum, are you an owner etc. (around the table)

GP(8am):

A(10am): L(10am): AA(10am): E(10am):

1. How would you define quality of care ?
2. Can you tell me about your experience with the QOF ?

Do you think QOF Is working as a quality incentive system? In what ways?

3. Has the QOF influenced/affected your professional practice? In what ways?
4. Has the QOF influenced /affected your patient management or handling?

Affected how you manage all aspects of the patient ?

5. Has the QOF influenced/affected the composition or mix of the patient population of your practice? More diabetic patients? More blood pressure patients? Less psychiatric patient and so on ?
6. Has the QOF influenced your or health professionals attitudes towards the patients or the treatment?
7. Has the QOF influenced cooperation fields or cooperation partners (nurses, psychologists and so on) ?
8. Has the QOF influenced/affected the professional development?
9. Has the QOF made the everyday as GP better?
10. Do you encounter any *ethical dilemmas due to QOF* ?
11. *Should the percentage* share of QOF bin bigger ? Smaller?
12. What do you think would be the best way to improve quality in general practice?
13. Do you have anything you want to add?

Appendix 3

UiO : **Faculty of Medicine**
University of Oslo
Date: December 5th 2012
Your ref.:
Our ref.:

Information about the project

The aim of the conversation is to learn from your experiences with the Quality and Outcomes Framework in the general practice.

I will write a master thesis in Norwegian focusing on the potential role of payment for performance as an instrument to improve quality in general practice. Material from interviews with GPs will be part of the basis for writing this thesis.

My work is supervised by medical doctor and professor Jan Frich, University of Oslo.

No name or practice location will be mentioned in the thesis, and I will guarantee that anonymity will be fully secured.

Sincerely yours,

Kari Jussie Lønning
GP/master's student

Contact:
Head of Unit for Primary Care and Psychiatry
Department of professional affairs
Norwegian Medical Association
Postbox 1152 Sentrum
N-0107 Oslo
Norway
e-mail: kari.jussie.lonning@legeforeningen.no



Institute of Health and Society
Postal addr.: PO Box 1130 Blindern, 0318
Oslo
Visiting addr.: Stjerneblokk, Nedre Ullevål
9, 0850 Oslo

Phone: (+47) 22 85 05 50
Telefax: (+47) 22 85 05 90
postmottak@medisin.uio.no
www.med.uio.no/helsam
Org. no.: 971 035 854